

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:49:54 ; Search time 4459 Seconds
(without alignments)
12395.003 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaagcaacgt.....gaatttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|----------|--------------------|
| 1 | 1452 | 100.0 | 1602 | 3 | CNS0A0KS | BX833171 Arabidops |
| 2 | 1450 | 99.9 | 1656 | 3 | CNS0A1A4 | BX829836 Arabidops |
| 3 | 1446.8 | 99.6 | 1585 | 3 | CNS0A1A4 | BX829783 Arabidops |
| 4 | 606.8 | 41.8 | 624 | 1 | AV823628 | AV823628 AV823628 |
| 5 | 599.6 | 41.3 | 606 | 6 | CB261053 | CB261053 35-E9570- |
| 6 | 406.4 | 28.0 | 920 | 7 | CF215873 | CF215873 CAST0002 |
| 7 | 391 | 26.9 | 894 | 7 | CO117860 | CO117860 GR_Eb01O |
| 8 | 371 | 25.6 | 866 | 7 | CO095556 | CO095556 GR_Ea18E |
| 9 | 364.6 | 25.1 | 749 | 7 | CN187799 | CN187799 UCRC505_0 |
| 10 | 357.2 | 24.6 | 723 | 7 | CK319497 | CK319497 X9P11dhl |
| 11 | 355.2 | 24.5 | 825 | 7 | CO096428 | CO096428 GR_Ea19K |
| 12 | 347 | 23.9 | 348 | 7 | Z37245 | Z37245 ATTS3988 St |
| 13 | 343.2 | 23.6 | 2173 | 3 | AY596550 | AY596550 Saccharum |
| 14 | 339.8 | 23.4 | 742 | 7 | CV197356 | CV197356 CGF100395 |
| 15 | 338.4 | 23.3 | 2189 | 3 | AY105375 | AY105375 Zea mays |
| 16 | 330.2 | 22.7 | 977 | 7 | CK283827 | CK283827 EST746549 |
| 17 | 328.6 | 22.6 | 886 | 7 | CK289726 | CK289726 EST752448 |
| 18 | 325.8 | 22.4 | 849 | 7 | CK264004 | CK264004 EST710082 |
| 19 | 324 | 22.3 | 883 | 7 | CO073048 | CO073048 GR_Ea32I |
| 20 | 323.8 | 22.3 | 681 | 5 | BUB36279 | BUB36279 T08G02 P |
| 21 | 315.2 | 21.7 | 608 | 7 | CN905453 | CN905453 010928ABD |
| 22 | 313.6 | 21.6 | 613 | 7 | CN913815 | CN913815 030108ABM |
| 23 | 313.6 | 21.6 | 636 | 7 | CN913251 | CN913251 030107ABM |
| 24 | 312 | 21.5 | 663 | 1 | AJ805795 | AJ805795 AJ805795 |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 25 | 311.8 | 21.5 | 763 | 2 | BE034121 | BE034121 MG05C04 M |
| 26 | 307.2 | 21.2 | 594 | 7 | CN912222 | CN912222 021217ABM |
| 27 | 307.2 | 21.2 | 626 | 7 | CN862200 | CN862200 000818AAL |
| 28 | 305.2 | 21.0 | 721 | 6 | CB917703 | CB917703 VVD019F01 |
| 29 | 301.2 | 20.7 | 586 | 1 | AI998258 | AI998258 701544952 |
| 30 | 298.2 | 20.5 | 688 | 7 | CN884491 | CN884491 010811AAS |
| 31 | 294.8 | 20.3 | 717 | 7 | CF808960 | CF808960 P8HB04LXE |
| 32 | 294.8 | 20.3 | 764 | 5 | BQ045892 | BQ045892 EST595010 |
| 33 | 293 | 20.2 | 494 | 1 | AV784632 | AV784632 AV784632 |
| 34 | 292.8 | 20.2 | 796 | 7 | CK261683 | CK261683 EST707761 |
| 35 | 290 | 20.0 | 290 | 1 | AV525150 | AV525150 AV525150 |
| 36 | 283.8 | 19.5 | 536 | 5 | BQ588266 | BQ588266 S013218W- |
| 37 | 281.8 | 19.4 | 559 | 4 | BI471389 | BI471389 sag19h04. |
| 38 | 281.2 | 19.4 | 497 | 7 | CN865160 | CN865160 001001AAL |
| 39 | 280.8 | 19.3 | 500 | 2 | AW201710 | AW201710 sf06d03.Y |
| 40 | 280.4 | 19.3 | 600 | 4 | BG139883 | BG139883 EST480325 |
| 41 | 280.2 | 19.3 | 504 | 1 | AI988515 | AI988515 sd03d02.Y |
| 42 | 280.2 | 19.3 | 518 | 4 | BG882882 | BG882882 sae78h02. |
| 43 | 280.2 | 19.3 | 542 | 5 | BQ080012 | BQ080012 san04f04. |
| 44 | 280.2 | 19.3 | 849 | 7 | CV287071 | CV287071 68153.1 A |
| 45 | 279.2 | 19.2 | 583 | 2 | AW686368 | AW686368 NF040G09N |

ALIGNMENTS

RESULT 1
CNS0A0KS
LOCUS
DEFINITION
CNS0A0KS 1602 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from Clone
GSITSL37ZE02 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).
ACCESSION
BX833171.1 GI:42457670
VERSION
HTC; GSI.T.CDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
REFERENCE
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpetti,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat.M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1602)
AUTHORS
Genoscope.
TITLE
Submitted (19-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)
COMMENT
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/egbi-bin/ggb?source=Arabidopsis.
length
http://www.genoscope.cns.fr/egbi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1..1602
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

FEATURES
source

| /clone="GSLTSL1372E02" | | | | | | | | | |
|---|-----|---|-----|--|--|--|--|--|--|
| /tissue_type="Silique" | | | | | | | | | |
| /plasmid="pCWSFOR6" | | | | | | | | | |
| complement(1..1602) | | | | | | | | | |
| /gene="At5g54310" | | | | | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 100.0%; Score 1452; DB 3; Length 1602; | | | | | | | | | |
| Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCGCCATAGAAAGATTCCTTCAA | 60 | | | | | | |
| Db | 75 | ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCGCCATAGAAAGATTCCTTCAA | 134 | | | | | | |
| Qy | 61 | GGGCTCTTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA | 120 | | | | | | |
| Db | 135 | GGGCTCTTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCACAGA | 194 | | | | | | |
| Qy | 121 | TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT | 180 | | | | | | |
| Db | 195 | TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT | 254 | | | | | | |
| Qy | 181 | CTCGGGGTACACATATCGAAGGTTCCGATCTGCCACTCTGGACACATGGCTCCCGAGCAG | 240 | | | | | | |
| Db | 255 | CTCGGGGTACACATATCGAAGGTTCCGATCTGCCACTCTGGACACATGGCTCCCGAGCAG | 314 | | | | | | |
| Qy | 241 | GTTCGATTATATACGTCAATGCGAAATGATAAGCAAAATAGTTACTTGGGAAGCGAGCTA | 300 | | | | | | |
| Db | 315 | GTTCGATTATATACGTCAATGCGAAATGATAAGCAAAATAGTTACTTGGGAAGCGAGCTA | 374 | | | | | | |
| Qy | 301 | CCCCCAAACTATGATAGGTTGGATTCGAAATTTTATACGTGCAAAATATGAGAGAG | 360 | | | | | | |
| Db | 375 | CCCCCAAACTATGATAGGTTGGATTCGAAATTTTATACGTGCAAAATATGAGAGAG | 434 | | | | | | |
| Qy | 361 | AGATGGGTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGACAGCAAGCGGG | 420 | | | | | | |
| Db | 435 | AGATGGGTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGACAGCAAGCGGG | 494 | | | | | | |
| Qy | 421 | AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGGAATAGTATGATGCTGTAAAT | 480 | | | | | | |
| Db | 495 | AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGGAATAGTATGATGCTGTAAAT | 554 | | | | | | |
| Qy | 481 | TGTTTGGAGGAGCAAAATCTTCCAGCATCTAGCAAGAAATTAATGTTGCTGCAAG | 540 | | | | | | |
| Db | 555 | TGTTTGGAGGAGCAAAATCTTCCAGCATCTAGCAAGAAATTAATGTTGCTGCAAG | 614 | | | | | | |
| Qy | 541 | AGAATAAATCTTCCCGTGCCTCCCAAGGACCCAGTCAAGTTATAAGCCACAGAGAAA | 600 | | | | | | |
| Db | 615 | AGAATAAATCTTCCCGTGCCTCCCAAGGACCCAGTCAAGTTATAAGCCACAGAGAAA | 674 | | | | | | |
| Qy | 601 | ATGGAGTCTGACGTACTCTCAGTAGAGAGGAGAAACAGCGATTAATGTTGACACGCA | 660 | | | | | | |
| Db | 675 | ATGGAGTCTGACGTACTCTCAGTAGAGAGGAGAAACAGCGATTAATGTTGACACGCA | 734 | | | | | | |
| Qy | 661 | TCAGATCTCCAAAGGTGATTTCTGCTAGTATCTGTTTAAACATCTCAATGATGAT | 720 | | | | | | |
| Db | 735 | TCAGATCTCCAAAGGTGATTTCTGCTAGTATCTGTTTAAACATCTCAATGATGAT | 794 | | | | | | |
| Qy | 721 | TCGACTACAAATACCTCAGAGGCAACTCTGCGGATACCTCTGCGGATGATAACTCATGG | 780 | | | | | | |
| Db | 795 | TCGACTACAAATACCTCAGAGGCAACTCTGCGGATACCTCTGCGGATGATAACTCATGG | 854 | | | | | | |
| Qy | 781 | GCTGCTTTCACTGCTGGAAGTGGTCAACCGCAGAGAAAATTTGTCAAGCCCAAGCCT | 840 | | | | | | |
| Db | 855 | GCTGCTTTCACTGCTGGAAGTGGTCAACCGCAGAGAAAATTTGTCAAGCCCAAGCCT | 914 | | | | | | |
| Qy | 841 | GCTGAGACAGTCTCTCCAGCTTCATCTCTGATCTTGGATTTCTTTAAGACACA | 900 | | | | | | |
| Db | 915 | GCTGAGACAGTCTCTCCAGCTTCATCTCTGATCTTGGATTTCTTTAAGACACA | 974 | | | | | | |
| Qy | 901 | CCTAATTTAACTCAACACAGCACAAAGATGTGAAAGGCGATATCATGAGCCTGTTT | 960 | | | | | | |

| | | | |
|----|------|---|------|
| Db | 975 | CCTAATTTAACTCAACAGCACCAAAAGATGTGAAAGCGGATATCATGAGCCTGTTT | 1034 |
| Qy | 961 | GAGAAGACGAATATAGTATGCGCTTTTGCCATGATCAGCAACAGGTTGCTATGCTCGCT | 1020 |
| Db | 1035 | GAGAAGACGAATATAGTATGCGCTTTTGCCATGATCAGCAACAGGTTGCTATGCTCGCT | 1094 |
| Qy | 1021 | CAGCAGCAAGCCCTTTTACATGGCTGCGAGCGAAAGCTGCTGGAGGCATCTCCAAAAGCGCGT | 1080 |
| Db | 1095 | CAGCAGCAAGCCCTTTTACATGGCTGCGAGCGAAAGCTGCTGGAGGCATCTCCAAAAGCGCGT | 1154 |
| Qy | 1081 | AATCAACAAAGCTATTGCTTAATGCTCTTTAAAGTCTTCTGCAAAATGCTCAAAAGCGCGC | 1140 |
| Db | 1155 | AATCAACAAAGCTATTGCTTAATGCTCTTTAAAGTCTTCTGCAAAATGCTCAAAAGCGCGC | 1214 |
| Qy | 1141 | GGCTACAGATCCCGGAATGACTTAACCCGTAAGTGGTCAAGCTGATCTCCAGAACTT | 1200 |
| Db | 1215 | GGCTACAGATCCCGGAATGACTTAACCCGTAAGTGGTCAAGCTGATCTCCAGAACTT | 1274 |
| Qy | 1201 | ATGCAAAACATGAATATGAACCGCAACATGAACACAGAGACCCGCAACCCGCAAGAGAAC | 1260 |
| Db | 1275 | ATGCAAAACATGAATATGAACCGCAACATGAACACAGAGACCCGCAACCCGCAAGAGAAC | 1334 |
| Qy | 1261 | ACTCAATAATCCCATCATCAGTTTCTTACAAATGGGTCAAGTAAATCAAGTGAACGGT | 1320 |
| Db | 1335 | ACTCAATAATCCCATCATCAGTTTCTTACAAATGGGTCAAGTAAATCAAGTGAACGGT | 1394 |
| Qy | 1321 | ATGACCCCAACCTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAAGAGCACC | 1380 |
| Db | 1395 | ATGACCCCAACCTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAAGAGCACC | 1454 |
| Qy | 1381 | ACACCATCTTCACAATCAGCAAAAGACTTTGATTTCTTCTTCTGATGGAATGTTT | 1440 |
| Db | 1455 | ACACCATCTTCACAATCAGCAAAAGACTTTGATTTCTTCTTCTGATGGAATGTTT | 1514 |
| Qy | 1441 | ACAAAACATTGA 1452 | |
| Db | 1515 | ACAAAACATTGA 1526 | |

| | | | | | |
|------------|---|-------------|------|--------|-----------------|
| RESULT 2 | CNSO1A14 | 1656 bp | mRNA | linear | HTC 06-FEB-2004 |
| LOCUS | Arabidopsis thaliana Full-length cDNA Complete sequence from clone | | | | |
| DEFINITION | GS1T3B3Z12 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress). | | | | |
| ACCESSION | BX829836 | | | | |
| VERSION | BX829836.1 | GI:42458516 | | | |
| KEYWORDS | HTC; GS1T cDNA. | | | | |
| SOURCE | Arabidopsis thaliana (thale cress) | | | | |
| ORGANISM | Arabidopsis thaliana | | | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. | | | | |
| AUTHORS | Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M. | | | | |
| TITLE | Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 1656) | | | | |
| AUTHORS | Genoscope. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr - Web : www.genoscope.cns.fr) | | | | |
| COMMENT | The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. | | | | |

URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 Genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>.

FEATURES

source
 1..1656
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTFB3ZH12"
 /tissue_type="Flowers and buds"
 /plasmid="pCMVSPORT6"
 complement(1..1656)
 /gene="At5g54310"

ORIGIN

Query Match 99.9%; Score 1450; DB 3; Length 1656;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAACGAGAAAGCAACGCTCTAAGGAGCTTAATGCCCGCCATAGAAAAGATTCTTGAAGG 62
 DB 1 GAACGAGAAAGCAACGCTCTAAGGAGCTTAATGCCCGCCATAGAAAAGATTCTTGAAGG 60
 QY 63 GCTTCTTAAACATCCAGAGAAACAGAGAATGTCTGACTGCAATGCAAAACAAAGGTCCTAAGATG 122
 DB 61 GCTTCTTAAACATCCAGAGAAACAGAGAATGTCTGACTGCAATGCAAAACAAAGGTCCTAAGATG 120
 QY 123 GCTAGTGTAAATAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGTCT 182
 DB 121 GCTAGTGTAAATAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGTCT 180
 QY 183 CGGGGTACACATATCGAAGGTTGCGATCTGCCACTCTGGACACATGGCTCCCCGAGCAGGT 242
 DB 181 CGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCCGAGCAGGT 240
 QY 243 TGCATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAAAGCAGACTACC 302
 DB 241 TGCATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAAAGCAGACTACC 300
 QY 303 CCCAACTATGATAGATTGGATTGAGNAATTTTATACGTGCAAAAGTATGAAGAGAG 362
 DB 301 CCCAACTATGATAGATTGGATTGAGNAATTTTATACGTGCAAAAGTATGAAGAGAG 360
 QY 363 ATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGCAGGAAACGGCGAA 422
 DB 361 ATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTAGAGTCGAGCAGGAAACGGCGAA 420
 QY 423 ATCTGTGGAGAGAGTGGCCGGGATATGAGCATGACATAGTAGTCTGTAAATTT 482
 DB 421 ATCTGTGGAGAGAGTGGCCGGGATATGAGCATGACATAGTAGTCTGTAAATTT 480
 QY 483 GTTTGAGGAGAGGAAAATATTTCCAGCATCTAGAACAGAAATATTTGCTGCAACGAG 542
 DB 481 GTTTGAGGAGAGGAAAATATTTCCAGCATCTAGAACAGAAATATTTGCTGCAACGAG 540
 QY 543 ATAAATCTTCCCGTCCCTCCCAAGGACCCAGTTCAGGTTTATAAGCCACACAGAAAT 602
 DB 541 ATAAATCTTCCCGTCCCTCCCAAGGACCCAGTTCAGGTTTATAAGCCACACAGAAAT 600
 QY 603 GGAGTCTGAGCTACTTCCAGTAGAGAGGGAGAAAACAGCAGTAATTTTGCACACGATC 662
 DB 601 GGAGTCTGAGCTACTTCCAGTAGAGAGGGAGAAAACAGCAGTAATTTTGCACACGATC 660
 QY 663 AGATCTCTCAAAAGGTGGATTTGCTACTGATCTGTTTAAATGCTATCAATGGATGATTC 722
 DB 661 AGATCTCTCAAAAGGTGGATTTGCTACTGATCTGTTTAAATGCTATCAATGGATGATTC 720
 QY 723 GACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCTGGCGATGATACTCATGGGC 782

DB 721 GACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCTGGCGATGATACTCATGGGC 780
 QY 783 TGGCTTTTCACTGCTCTGGAAGTGTCAAAAGGAGAGAAAATTTGTCACAGCCAGCCTGC 842
 DB 781 TGGCTTTTCACTGCTCTGGAAGTGTCAAAAGGAGAGAAAATTTGTCACAGCCAGCCTGC 840
 QY 843 TGAGAGCAGATTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC 902
 DB 841 TGAGAGCAGATTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC 900
 QY 903 TAATTTAACTCAACAGCACCACAAAGATGTGAAGGCGATATCATGAGCCTGTTGA 962
 DB 901 TAATTTAACTCAACAGCACCACAAAGATGTGAAGGCGATATCATGAGCCTGTTGA 960
 QY 963 GAACACGAATATAGTATGCGCTTTTGCATCATCAGCAACAGGTTGCTATGCTGCTCA 1022
 DB 961 GAACACGAATATAGTATGCGCTTTTGCATCATCAGCAACAGGTTGCTATGCTGCTCA 1020
 QY 1023 GCAGCAAGCCCTTTTACATGGCTGAGCGAAAGCTGTGGAGGCACTCCAAACGGCGTGA 1082
 DB 1021 GCAGCAAGCCCTTTTACATGGCTGAGCGAAAGCTGTGGAGGCACTCCAAACGGCGTGA 1080
 QY 1083 TCAACAGCTATTTGCTAATGCTCTTAAAGCTTCTGCAAAATTTGTTCAAAACCCCGCGG 1142
 DB 1081 TCAACAGCTATTTGCTAATGCTCTTAAAGCTTCTGCAAAATTTGTTCAAAACCCCGCGG 1140
 QY 1143 CTACAGATCCCCGGAATGACTAACCCGTAAGTGTCAAGCTGATCTCCAGAAACTTAT 1202
 DB 1141 CTACAGATCCCCGGAATGACTAACCCGTAAGTGTGTCAAGCTGATCTCCAGAAACTTAT 1200
 QY 1203 GCAAAACATGAATGAACGCAACATGAACGAGACCCGCAACACCGCAAGAGAACAC 1262
 DB 1201 GCAAAACATGAATGAACGCAACATGAACGAGACCCGCAACACCGCAAGAGAACAC 1260
 QY 1263 TCTCAATATCCATCATCTCAGTTTCTACAATGGTCAAGCTATCAAGTGAACGGTAT 1322
 DB 1261 TCTCAATATCCATCATCTCAGTTTCTACAATGGTCAAGCTATCAAGTGAACGGTAT 1320
 QY 1323 GACCCCAACTCAACCGGTAACCTCAGTCAATATCGCAACCCCAACCAACAGCACCAC 1382
 DB 1321 GACCCCAACTCAACCGGTAACCTCAGTCAATATCGCAACCCCAACCAACAGCACCAC 1380
 QY 1383 ACCATCTTCAATCAGGCAAGACTTTGATTCTCTCTTGATGGATGGAATGTTTAC 1442
 DB 1381 ACCATCTTCAATCAGGCAAGACTTTGATTCTCTCTTGATGGATGGAATGTTTAC 1440
 QY 1443 AAAACATGA 1452
 DB 1441 AAAACATGA 1450

RESULT 3

CNS0A1AM 1585 bp mRNA linear HTC 06-FEB-2004
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSTFB3ZH03 of Flowers and buds of strain col-0 of Arabidopsis
 thaliana (thale cress).
 BX829783
 BX829783.1 GI:42458494
 HTC; GSLT cDNA.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 1585)
 Castelli V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Searpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation

LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCES
 AUTHORS
 TITLE

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1585)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/banque_Projet_EF/Pull length
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
 FEATURES
 Location/Qualifiers
 1..1585
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="CSLFB36ZH03"
 /issue_type="Flowers and buds"
 /plasmid="pCMVSPORT 6"
 complement(1..1585)
 /gene="At5g54310"
 gene
 ORIGIN
 Query Match 99.8%; Score 1446.8; DB 3; Length 1585;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1448; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 GAACGAGAAGCCACGCTCTTAAGAGGCTTAATGCGCCCATAGAAAAGATTCTTGAAGG 62
 Db 1 GAACGAGAAGCCACGCTCTTAAGAGGCTTAATGCGCCCATAGAAAAGATTCTTGAAGG 60
 QY 63 GCTTCTTAACATCCAGAGAACAGAGATGTGTGATCGCACTGCAAAACAAAGGTCCAGATG 122
 Db 61 GCTTCTTAACATCCAGAGAACAGAGATGTGTGATCGCACTGCAAAACAAAGGTCCAGATG 120
 QY 123 GCGTAGTTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCAGAGATCT 182
 Db 121 GCGTAGTTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCAGAGATCT 180
 QY 183 CGGGGTACACATATCGAAGGTTGATCTGCACTCTGGACACATGCTCCCGAGCAGGT 242
 Db 181 CGGGGTACACATATCGAAGGTTGATCTGCACTCTGGACACATGCTCCCGAGCAGGT 240
 QY 243 TGCATTTATACAGTCAATGGAAATGATAAGCAATAGTTACTGGGAAGCAGAGTACC 302
 Db 241 TGCATTTATACAGTCAATGGAAATGATAAGCAATAGTTACTGGGAAGCAGAGTACC 300
 QY 303 CCCAACTATGATAGATTGGAAATTCAGAAATTTATACGTGCAGATATGAGAGAGAG 362
 Db 301 CCCAACTATGATAGATTGGAAATTCAGAAATTTATACGTGCAGATATGAGAGAGAG 360
 QY 363 ATGGGTTTCTAGAGGGGAAAGGCTAGATCAGCTCTCTAGAGTCGAGCAGGAACGGCGAA 422
 Db 361 ATGGGTTTCTAGAGGGGAAAGGCTAGATCAGCTCTCTAGAGTCGAGCAGGAACGGCGAA 420
 QY 423 ATCTGTGAGAGAAAGTGGGCGGGATATGAGCATATGAGCATATGAGTATGCTGTAAATTT 482
 Db 421 ATCTGTGAGAGAAAGTGGGCGGGATATGAGCATATGAGCATATGAGTATGCTGTAAATTT 480
 QY 483 GTTTGAGGAGAGAAAACCTATTTCCAGCATCTAGAACAGAAATATGTTGCTGCAACGAG 542
 Db 481 GTTTGAGGAGAGAAAACCTATTTCCAGCATCTAGAACAGAAATATGTTGCTGCAACGAG 540

QY 543 AATAAATCTTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAAAAT 602
 Db 541 AATAAATCTTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAAAAT 600
 QY 603 GGAGTCTGCAGTACTCCAGTAGAGAGGGAGAAAACAGCAGTAATGTTGCACAGCATC 662
 Db 601 GGAGTCTGCAGTACTCCAGTAGAGAGGGAGAAAACAGCAGTAATGTTGCACAGCATC 660
 QY 663 AGATCTCCAAAGGTGGATTTTGTACTGATCTGTTTAAACATGCTATCAATGGATGATTC 722
 Db 661 AGATCTCCAAAGGTGGATTTTGTACTGATCTGTTTAAACATGCTATCAATGGATGATTC 720
 QY 723 GACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCCGATATCAATCTATGGC 782
 Db 721 GACTACAAATACCTCAGAGGCAACTCTCTGGCGATCTCTCCGATATCAATCTATGGC 780
 QY 783 TGGCTTTTCAGTCTCTGGAAGTGTCAACCGGAGAGAAAATTTGTCACAGCCACCTGCG 842
 Db 781 TGGCTTTTCAGTCTCTGGAAGTGTCAACCGGAGAGAAAATTTGTCACAGCCACCTGCG 840
 QY 843 TGAGACGAGTTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC 902
 Db 841 TGAGACGAGTTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTTAAAGGACACACC 900
 QY 903 TAATTTAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTTGA 962
 Db 901 TAATTTAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTTGA 960
 QY 963 GAAACGATATAGTATGCGCTTTTGGCATGATCAGCAACAGTTGCTGCTCGCTCA 1022
 Db 961 GAAACGATATAGTATGCGCTTTTGGCATGATCAGCAACAGTTGCTGCTCGCTCA 1020
 QY 1023 GCAGCAGCCCTTTACATGGCTGCAGGGAAGCTGCTGGAGGCACTCCAAACGGCGTGAA 1082
 Db 1021 GCAGCAGCCCTTTACATGGCTGCAGGGAAGCTGCTGGAGGCACTCCAAACGGCGTGAA 1080
 QY 1083 TCAACAGCTATGCTAATGCTCTTAACGTAGCTTTGCAAAATTTGGTCAAAACCCCGCGG 1142
 Db 1081 TCAACAGCTATGCTAATGCTCTTAACGTAGCTTTGCAAAATTTGGTCAAAACCCCGCGG 1140
 QY 1143 CTACAGATCCCGGAATGACTTAACCGTAGTGGTCAAGCTGATCTCCGAAACTTAT 1202
 Db 1141 CTACAGATCCCGGAATGACTTAACCGTAGTGGTCAAGCTGATCTCCGAAACTTAT 1200
 QY 1203 GCAAAACATGATATGAACGCAAAACATGAACACGAGACCCGCAACCCGCAAGAGAACAC 1262
 Db 1201 GCAAAACATGATATGAACGCAAAACATGAACACGAGACCCGCAACCCGCAAGAGAACAC 1260
 QY 1263 TCTACAATACCCATCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTAT 1322
 Db 1261 TCTACAATACCCATCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTAT 1320
 QY 1323 GACCCCAACTCAACCGGTAAACCTCAGTCAATATCGCAACCCCAACCAAGCACCAC 1382
 Db 1321 GACCCCAACTCAACCGGTAAACCTCAGTCAATATCGCAACCCCAACCAAGCACCAC 1380
 QY 1383 ACCATCTTCAATCAGGCAAGACTTTGATTTCTTCTTGTGATGGAATTTTCAAC 1442
 Db 1381 ACCATCTTCAATCAGGCAAGACTTTGATTTCTTCTTGTGATGGAATTTTCAAC 1440
 QY 1443 AAAACATTGA 1452
 Db 1441 AAAACATTGA 1450

RESULT 4
 AV823628
 LOCUS AV823628 RAPUS Arabidopsis thaliana cDNA clone RAPL05-21-P04 5',
 DEFINITION mRNA sequence.
 ACCESSION AV823628
 VERSION AV823628.1 GI:1986568
 KEYWORDS EST.

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 624)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayaishizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified paluascript vector as a SctI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.

FEATURES
source
1..624
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-21-P04"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/notes="Site 1: SctI; Site 2: XhoI; subjected to
dehydration-treated (1,2,5,10,24 hr)"

ORIGIN
Query Match 41.8%; Score 606.8; DB 1; Length 624;
Best Local Similarity 99.3%; Pred. No. 1.7e-166;
Matches 608; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAACCAACGCTCTTAAGAGAGCTTAATGCCCGCCATGAGAAAGATTTCTTGAA 60
DB 13 ATGACGAGGAAAGCCAAAGCTCTCTAAGAGCTTAATGCCCGCCATGAGAAAGATTTCTTGAA 72
QY 61 GGCTTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCCAAG 120
DB 73 GGCTTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCCAAG 132
QY 121 TGGCTAGTGTAAATTTAGTATCTTTATCTGCAATGTTCTGGATTCACAGAGT 180
DB 133 TGGCTAGTGTAAATTTAGTATCTTTATCTGCAATGTTCTGGATTCACAGAGT 192
QY 181 CTGGGGTACACATATCGAAGTTTCATCTGCCACTCTGGACATCGGCTCCCGGAGCAG 240
DB 193 CTGGGGTACACATATCGAAGTTTCATCTGCCACTCTGGACATCGGCTCCCGGAGCAG 252
QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 253 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 312
QY 301 CCCCCAACTATCATAGAGTTGGAAATGAGAAATTTATAGTCCCAAGTATGAGAGAG 360
DB 313 CCCCCAACTATCATAGAGTTGGAAATGAGAAATTTATAGTCCCAAGTATGAGAGAG 372
QY 361 AGATGGTTCTTAGGGGAAAAGGCTAGATCACTCTAGATCGCAGAGGAAAGCGCG 420
DB 373 AGATGGTTCTTAGGGGAAAAGGCTAGATCACTCTAGATCGCAGAGGAAAGCGCG 432
QY 421 AAATCTGTGAGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 480
DB 433 AAATCTGTGAGAGAGAGTGGCGCGGATATGAGCATGGACATAGTAGTCTCTGTAAT 492

QY 481 TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGACAAAGAAATAATTTGCTGCAACG 540
DB 493 TTGTTTGAGGAGAGAAACTATTTCAGCATCTAGACAAAGAAATAATTTGCTGCAACG 552
QY 541 AGAATAAATCTTCCGTCCTCCCTCCCAAGACCACTCAGGTTTATAAGCCACAGCAGAA 600
DB 553 AGAATAAATCTTCCGTCCTCCCTCCCAAGACCACTCAGGTTTATAAGCCACAGCAGAA 612
QY 601 ATGGAGTCTGCA 612
DB 613 ATGGAGTCTGCA 624

RESULT 5
CB261053
LOCUS
DEFINITION
35-E9570-012-004-E09-T7R MP1Z-ADIS-012 Arabidopsis thaliana cDNA
clone MP1Zp769E094Q 5-PRIME, mRNA sequence.
CB261053
CB261053.1 GI:32885826
EST
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
Arabidopsids
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 606)
Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T.,
Mitchell-Olds, T. and Weissshaar, B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)
22683290
12799357
Contact: Weissshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weissshaar@mpiz-koeln.mpg.de
Insert Length: 606 Std Error: 0.00
Plate: 4 row: E column: 09
Seq primer: T7R; CTAATACGACTCTACTATAGGA.
Location/Qualifiers
1..606
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Landsberg erecta"
/db_xref="GABI:589856"
/db_xref="taxon:3702"
/clone="MP1Zp769E094Q"
/tissue_type="whole plant"
/dev_stage="adult plant, mixed stresses"
/lab_host="E. coli XL1-Blue MRF"
/clone_lib="MP1Z-ADIS-012"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Landsberg
erecta; six weeks old total plants grown under long-day
conditions in soil, whole adult plants were treated for 24
hours with different stresses, (1) at 4M-0 C in the dark,
(2), at 37 Grad C in the dark, (3) lying in the lab after
removing from soil, (4) in the greenhouse after wounding
leaves with a forceps, (5) in the lab watering with a 150
mM NaCl solution, (6) at 26 Grad C in the light/UV; equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCACGCTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd


```

593 AATTCATTGAGGAAAAGAAGATGTTTCAGGCACCTAGTATATAAAGATAAACGGGCGCTGCT 652
538 ACAGAGTAATATCTCCCTGCTCCCGAGGACCCAGTCAGG-----TTATTAAG 588
653 ACAAGGATAGGCTTCTGCTGCTCTTAAGAGACCTGAGCCAGTTGCCCTATTCCAAG 712
589 CCACAG-----CAGAAATGAGTCTGCGAGTACTCCAGTAGAGAGGAGAAACAA 639
713 CTCTCATAGTTACTCAGAAACAGAAACCATCAGTGGCACAAGCTGAATCAGTAAAGCAG 772
640 GCAGTAATATGTCACAGATCAGATCTCCAAAGGTGGATTTGCTACTGATCTGTTT 699
773 GCTGCAGATCTACTCCAGTAGTTCTTCCCTCCCTAAAGTTGATTTGCTATTGATCTTTT 832
700 AACATGCTATCATGATGATTCGACTACATATACCTCAGAGCAACTCTGGCGATACT 759
833 TACATGCTTTCTATGGATGATCTTACTGAAATGCGCTCAGAGCA-----GCC 880
760 CTGCGCGATGATACTCATGGGCTGCTTTCAGTCTGCTG 799
881 TCTGCTGATGATAATGCAATGGGCTGTTTCCAGTCTGCTG 920

COL17860      894 bp mRNA linear EST 16-JUN-2004
GR_Eb01018.f GR_Eb Gossypium raimondii cDNA clone GR_Eb01018 5',
mRNA sequence.
COL17860      1 GI:48816547
COL17860      1
EST.
Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 894)
Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Stum.D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 01 row: O column: 18.
Location/Qualifiers
1. .894
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb01018"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional clones into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
Query Match 26.9%; Score 391; DB 7; Length 894;
Best Local Similarity 71.0%; Pred. No. 5e-103;
Matches 562; Conservative 0; Mismatches 220; Indels 9; Gaps 3;

QY 1 ATGAACGAGAAAGCCACGCTCTTAAGAGAGCTTAATGCCCGCATAGAAAGATTCTTGAA 60
|||||
DB 2 ATGAACGAGAGGCCAGCGTTACCAAGCTGTCTACAGCCAGACACCGAAAGATTATTGAA 61
|||||

```

```

61 GGGCTTCTTAAACATCCAGAGAACAGAGAATGTGCTGACTGCACAAACAAAGAGSTCCAAGA 120
|||||
62 AGTCTTCTTAAATTTGCCCGAGATAGGGAATGCGCGACTGCAAGCCAAAGAGSTCCGAGA 121
|||||
121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
|||||
122 TGGGCAAGTGTGAATTTAGGTATCTTTATATGCAATGCAATGTTCTGGGATTCACAGGAGT 181
|||||
181 CTCGGGGTACACATATCGAAGGTTCCGATCTGCACTCTGGACACATGGCTCCCGAGCAG 240
|||||
182 CTGCGGGTACACATATCGAAGGTTCCGATCTGCTACATCTAGACACATGGCTTCTGAGCAG 241
|||||
241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
|||||
242 GTTGCTTTTATTCATCAATGGGAATGAAAAGGCAACACAGTTACTTGGGAAGCTGAGTTA 301
|||||
301 CCCCCAAATATGATAGAGTTGGAATTTGAGAAATTTTATCTGTCGAAAGTATGAAGAGAG 360
|||||
302 CCCCCAAATATGATAGAGTTGGAATTTGAGAAATTTTATCTGTCGAAAGTATGAAGAGAG 361
|||||
361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAACGG--- 417
|||||
362 AGATGGGTTCTAGAGATGGGAATTCAAATCACCACCTAGAGAGGTTGGACGAAGGGCT 421
|||||
418 CGGAATCTGTGGAGAGAAAGTGGGCGGATATGAGCATGAGCATAGTAGTAGTCTCTGTA 477
|||||
422 CCTCACAATTGGCAGAGACCTAATGAACAAGTGGTCAATGGGCACATAGTAATTCGAG 481
|||||
478 AATTTGTTGGAGAGAGAAACATTAATTCAGCATCTAGAACAAAGAAATTAATGTTGCTGA 537
|||||
482 AATTCATTTGAGGAAGAGGAGAAACAACAAGCACTTGGTCAGAAAAGAAATCTTCTCTGA 541
|||||
538 ACGAGAAATAATCTTCCCGTGCCTCCCAAGGACCCAGTCAGGTTTATAAGCCACAGCAG 597
|||||
542 ACAAGAGTTAGTCTTCTCTGTTCTCTTAAGGACCTGATCAGGTTACA---CCCGTACAG 598
|||||
598 AAAATGAGTCTGACAGTACTCTCAGTAGAGGGGAGAAACAAGCAGTAATATGTTGCCA 657
|||||
599 AAGCCTGAACCACTGTTTGGCCAGCTGAGGCAACAAGCCAGCTGTAGAGACTGCTCCC 658
|||||
658 GCATCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAACTGCTATCAATGGAT 717
|||||
659 GTAGCCATCGCTCTTAAGAGTCGATTTATGCTACAGACCTTTTCAACATGCTGCTTGGAT 718
|||||
718 GA---TTGCACTACAAATACCTCAGAGCAACTCTCGCGGATACTCTCCCGATGATAAC 774
|||||
719 GATGTCCAAGTGAGATGTTCTGAGGCAACTTCTACTGATGTTGGCAGGCTCCAG 778
|||||
775 TCATGGGCTGG 785
779 TCTGCTGCTGG 789

RESULT 8
CO095556
LOCUS
DEFINITION
GR_Eb01018.r GR_Ea Gossypium raimondii cDNA clone GR_Eb01018 3',
mRNA sequence.
CO095556
CO095556.1 GI:48794242
VERSION
EST.
KEYWORDS
SOURCE
Gossypium raimondii
Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 866)
Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Stum.D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
TITLE
JOURNAL
COMMENT

```

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 18 row: B column: 01.
Location/Qualifiers
1. .866
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eal8E01"
/tissue_type="whole seedlings"
/dev_stages="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

FEATURES
source

Query Match 25.6%; Score 371; DB 7; Length 866;
Best Local Similarity 73.1%; Pred. No. 3.8e-97;
Matches 505; Conservative 0; Mismatches 180; Indels 6; Gaps 2;

ORIGIN

QY 1 ATGACGAGAGGCAACGCTCTTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGA 60
DB |||||
QY 179 ATGAACGAGAGGCGAGCGTTACCAAGGAGCTCAACGCCAGACCCGAAAGATTATTGA 238
DB |||||
QY 61 GGGCTCTTTAAACATCCAGAGACAGAGAATGTGCTGACTGCAAAACAAAGAGTCCAAGA 120
DB |||||
QY 239 AGTCTCTTAATTCGCGGAGATAGGATGCGCGACTGCAAGCCAAAGGTCGCGA 298
DB |||||
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGAGT 180
DB |||||
QY 299 TGGGCAAGTGTGAATTTAGGTATCTTTATGATGATGCAATGTTCTGGGATTCACAGAGT 358
DB |||||
QY 181 CTCGGGTACACATATGAGGTTGATGCTGCCACTCTGGACATGCTCCCGGACG 240
DB |||||
QY 359 CTGGGGTACACATATGAGGTTGATGCTGCTACATAGACATGCTTCCTGGAGCAG 418
DB |||||
QY 241 GTTGCAATTATACAGTCAATGGGAATGATAAGCAAAATGTTACTGGGAAGCAGAGCTA 300
DB |||||
QY 419 GTTGCTTTTATTCATCAATGGGATGATAAGGCAACAGTACTGGAGCTGAGTTA 478
DB |||||
QY 301 CCCCCAACTATGATAGAGTTGGAATGAGAAATTTATACGTGCAAAAGTATGAAGAGAAG 360
DB |||||
QY 479 CCCCCCTAATATGATAGAGTTGGAATTCAGAACTTCATCCGTGCAAAAGTATGAGGAAAA 538
DB |||||
QY 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACTCTAGATGCTGAGCAGGAGGCGGG 420
DB |||||
QY 539 AGATGGGTTCTAGAGATGGAATTCAAATCACCACCTGAGAGGTTGGACGAAAGGCT 598
DB |||||
QY 421 AAATC---TGTGGAGAGAGTGGGCGGATATGAGCATGATGATAGTGTAGTCTGTA 477
DB |||||
QY 599 CCCTCAGATGGCAGAGACCTAATGAAACAGTGTCTATGGGCACATTAATGATTCGAG 658
DB |||||
QY 478 AATTGTTTGGAGGAGGAAAATCTATTCAGCATCTAGAACAAAGAAAATATGTTGCTGCA 537
DB |||||
QY 659 AATTCATTTGAGGAAGGAGGACAAACAAAGCATTGCTCAGAAAGAAAATCTTCCTGCA 718
DB |||||
QY 538 ACGAGAAATAATCTTCCGCTGCTCCCAAGACCCAGTCAAGTTATTAAGCCACAGCAG 597
DB |||||
QY 719 ACAAGAGTTAGTCTTCTGTTCTCTCTCAAGGACCTGATCAGGTTACA---CCGCTACAG 775
DB |||||
QY 598 AAATGGAGTCTGACGACTCTCCAGTAGAGGGGAGAACAGCAGTAATGTTGCAACA 657
DB |||||
QY 776 AAGCCTGAACCAAGTTGTTGCGCAGCTGAGGCAACAAAGCAGCTGTGAGACTGCTCCC 835
DB |||||
QY 658 GCATCAGATCTCCAAAGGTTGATTTTCTTA 688

Db 836 GTAGCCATCGCTCTCTAAAGTCGATTATGCTA 866

RESULT 9
CN187799
LOCUS
DEFINITION
CN187799
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CN187799 749 bp mRNA linear EST 15-APR-2004
UCRCS05_0005A10_r Washington Navel Orange stored Fruit Pulp cDNA
Library Citrus sinensis cDNA clone CS_WEB0005A10, mRNA sequence.
CN187799
GI:46212736
EST.
Citrus sinensis
Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 749)
Close,T.J., Rose,M.L., Arpaia,M.L., Federici,C.F., Fenton,R.D.,
Manamaker,S., Focht,E., Sievert,J., Robinson,P., Kim,H.R.,
Kudrna,D., Stum,D., Yost,D. and Wing,R.
Development of EST Resources and New Genetic Markers for California
Citrus - Washington Navel Orange Stored Fruit Pulp
Unpublished (2004)
Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
Tel: 9097873318
Fax: 9097874437
Email: timothy.close@ucr.edu
Seq primer: T3.
Location/Qualifiers
1. .749
/organism="Citrus sinensis"
/mol_type="mRNA"
/cultiva="Washington navel"
/db_xref="taxon:2711"
/clone="CS_WEB0005A10"
/tissue_type="Pulp"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TJCL21"
/clone_lib="Washington Navel Orange Stored Fruit Pulp cDNA
Library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Fruits were collected
January-March 2003 (Federici, Rose lab; Focht, Sievert &
Robinson, Arpaia lab). Four samples related to storage
conditions were produced: 1) fresh-picked in Mentone
(Arnott Brothers Enterprises, Mentone, CA), 2) after 21
days storage at 5C at Kearney then transported to UC
Riverside on ice, 3) after 5 additional days storage at
11C at Kearney, sampled immediately, 4) fruit grown in
southern CA were obtained from Redlands Football Packing
House after commercial packing, X-ray irradiated at 300 Gy
by Surebeam, then stored 1 day at ambient temperature.
Pulp tissue (juice vesicles) were collected. Tissues were
snap frozen and then stored at -80C until further
processing. Fenton (Close lab) purified RNA by the phenol
method described in J. Japanese Soc. Hort. Sci. 1996. 64
(4): 809-814, purified poly(A) mRNA using a PolyAtrack
mRNA Isolation System IV (Promega), produced a primary
cDNA library using a lambda ZAP XR cDNA Synthesis Kit
(Stratagene), then mass-excised one million pfu from the
primary library to produce a phagemid population.
Phagemids were plated, plasmid DNA purified, cDNA clones
archived, and DNA sequences determined bi-directionally
using an ABI3730 at the Arizona Genomics Institute,
University of Arizona (Kim, Kudrna, Stum, Yost, Wing).
Chromatogram files were downloaded by FTP to UC Riverside
(by Close), then processed at UC Riverside (by Manamaker,
Close lab) using the HarVEST pipeline
(<http://harvest.ucr.edu>) to remove vector and cloning
oligo sequences and various contaminants, and to trim to

high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

| | | | | | | |
|------------|-----|---|-----|---------------------------------------|---|-----|
| ORIGIN | | Query Match | | 24.6%; Score 357.2; DB 7; Length 723; | | |
| | | Best Local Similarity | | 71.0%; Pred. No. 4e-93; | | |
| | | Matches 508; Conservative | | 0; Mismatches 189; Indels 18; Gaps 2; | | |
| QY | 1 | ATCAACGAGAACCCACGCTCTTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAA | 60 | 19 | GTCTCTAAGGAGCTTAATGCCCGCATAGAAAGATTCTTGAAGGGCTTCTTAAACATCA | 78 |
| DB | 104 | ATGAACGAGAGGCCAACGCTCTTAAGAGCTTAACGCTAGACACAGGAAGATTCTGGAA | 163 | 4 | GTCTCTAAAGAGCTCAATGCTAGACACAGAAAGATCTTGAAGGCCCTCTCAATTTGCCT | 63 |
| QY | 61 | GGGCTCTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAAACAAAAGTCCAAGA | 120 | 79 | GAGAACAGAGAAATGTGCTGACTGCAAAAACAAAAGTCCAAGATGGGCTTAGTGTAAATTA | 138 |
| DB | 164 | GGGCTCTTTAAGTTGCCAGAGAAATAGGGAGTGGCTGACTGCAAAAGCCAAAGTCCAAGA | 223 | 64 | GAGATAGGGAATGTGCTGACTGCAAAAGTCCAAGTCCCAAGATGGGCAAGCGTGAATTTG | 123 |
| QY | 121 | TGGGCTAGTGTTAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT | 180 | 139 | GGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCG | 198 |
| DB | 224 | TGGGCTAGTGTGAATTTGGGAATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT | 283 | 124 | GGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATCTCA | 183 |
| QY | 181 | CTCGGGGTACACATATCGAAGGTTCCGATCTGCCACTCTGGAACACATGGCTCCCGGACAG | 240 | 199 | AAGGTTTCGATCTGCCACTCTGGAACACATGGCTCCCGGACAGGTTGCATTTATACAGTCA | 258 |
| DB | 284 | CTTGGGTATACATATCAAGGTTAGATCTGTACCTTGGATACCTGGCTTCGGAGCAG | 343 | 184 | AAGGTTTCGATCTGCCAACCCCTTGACACATGGCTTCAGAACACAGGTTGCATTTATTCATCA | 243 |
| QY | 241 | GTTCGATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA | 300 | 259 | ATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGTAGAGA | 318 |
| DB | 344 | GTTCGATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA | 403 | 244 | ATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGTTACCCCAAACTATGTAGAGA | 303 |
| QY | 301 | CCCCCAAACTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA | 360 | 319 | GTGGAATTTGGAATTTTATACGTCGCAAAAGTATGAAGAGAGAGATGGGTTCTTAGAGGG | 378 |
| DB | 404 | CTTCCAAACTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA | 463 | 304 | GTGGAATTTGGAATTTTATACGTCGCAAAAGTATGAAGAGAGAGATGGGTTCTTAGAGGG | 363 |
| QY | 361 | AGATGGGTTTCTAGAGGGGAAAGCTAGATCACTCTCTAGAG---TCGACAGGAGAACGG | 417 | 379 | GAAAAGGCTAGATCACTCTCTAGAGTCGACGAGGAGGCGGGAATCTCTGGAGAGAGT | 438 |
| DB | 464 | CGATGGGTTTCTAGAGATGGAACAAGCAAAATCACTCTCTAGAGGGCTGGAGGAGAGGCT | 523 | 364 | GGAAGAGCAACACCTCTCTCTAGTGGGCTGGATGAAAGATCTCTGCACT---CAGAGCCT | 420 |
| QY | 418 | CGGAATCTGTGAGAGAGTGGCGGGATATGACATGACATAGTAGTCTCTGTA | 477 | 439 | GGCGCGGATATGAGCATGACATAGTAGTCTCTCTGTAATTTGTTTGGAGAGAGGAA | 498 |
| DB | 524 | TCATTTCTGAGAGAGCTGGGGAAGAGAGTGGGCACTGTTACACTGATTAATTTCTGAA | 583 | 421 | GGTGAAGAGTGGGCATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG | 480 |
| QY | 478 | AATTTGTTGAGAGAGGAAACTATTTCCAGCATCTAGAACAAAGAAATATGTTGCTGCA | 537 | 499 | ACTATTCCAGCATCTAGAACAAAGAAATATGTTGCTGCAACGAGAAATATCTTCCGCTG | 558 |
| DB | 584 | AACTTATCTGAGGAAGGAGCATGTTCAAGCTCCAAGTACAAAGATAGTGTCTCTGCT | 643 | 481 | AATTCACAAAGTATCAAGCTCAAGGAACAGTGTCTGTCTACAGAAATAGGTGTTCTCTGTT | 540 |
| QY | 538 | ACGAGAAATATCTTCCGCTGCTCCCAAGGACCCAGTCAGGTTAT | 584 | 559 | CTTCCCAAGGACCCAGTCAGGTTATTAAGCCACAGCAGAAATGAGTCTGCAGCTACT | 618 |
| DB | 644 | GCAAGAAATGGTCTTCCACTTCCACTAGAGGACCTGATCAGGTTGT | 690 | 541 | CTCTCTAGAGACCCGAGCAGGTTACTCTCTCAAGAGCTCAACAGGTTGTTGAAAAGCT | 600 |
| RESULT 10 | | CK319497 | | 723 bp mRNA linear EST 11-MAY-2004 | | |
| LOCUS | | X911d11 Populus stem seasonal library Populus deltoides cDNA, mRNA | | sequence. | | |
| DEFINITION | | CK319497 | | EST. | | |
| ACCESSION | | CK319497.1 | | GI:47107920 | | |
| VERSION | | EST. | | | | |
| KEYWORDS | | Populus deltoides | | | | |
| SOURCE | | Populus deltoides | | | | |
| ORGANISM | | Populus deltoides | | | | |
| REFERENCE | | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | |
| AUTHORS | | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | | |
| TITLE | | rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus. | | | | |
| JOURNAL | | 1 (bases 1 to 723) | | | | |
| COMMENT | | Park, S. and Han, K.-H. | | | | |
| | | Gene expression profile during seasonal growth cycle in poplar tree | | | | |
| | | Unpublished (2003) | | | | |
| | | Contact: Kyung-Hwan Han | | | | |
| | | Department of Forestry | | | | |
| | | Michigan State University | | | | |
| | | 126 Natural Resources, East Lansing, MI 48824-1222, USA | | | | |
| | | Tel: 517 353 4751 | | | | |
| | | Fax: 517 432 1143 | | | | |
| | | Email: hanky@msu.edu. | | | | |
| RESULT 11 | | CO096428 | | 825 bp mRNA linear EST 16-JUN-2004 | | |
| LOCUS | | GR_Ea19K21.r GR_Ea Gossypium raimondii cDNA clone GR_Ea19K21.3, | | | | |
| DEFINITION | | mRNA sequence. | | | | |
| ACCESSION | | CO096428 | | | | |
| VERSION | | CO096428.1 | | GI:48795114 | | |
| KEYWORDS | | EST. | | | | |
| SOURCE | | Gossypium raimondii | | | | |

ORGANISM *Gossypium raimondii*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 825)
 Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.
 Global assembly of Cotton ESTs
 Unpublished (2004)
 Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: <http://genome.arizona.edu>
 Plate: 19 row: K column: 21.

FEATURES
 source 1..825
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_Eal9K21"
 /tissue_type="whole seedlings"
 /dev_stages="first true leaves"
 /lab_host="DH10B"
 /clone_libs="GR_Ea"
 /note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
 Query Match 24.5%; Score 355.2; DB 7; Length 825;
 Best Local Similarity 75.7%; Pred. No. 1.6e-92;
 Matches 454; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

Qy 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTATGCCC GCCCATAGAAAGATTCTTGAA 60
 Db 185 ATGAACGAGAAAGCCAGCGTTTACCAAGGAGCTCAACGCCAGACACCCGAAGATTATTGAA 244
 Qy 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAACAAACAAAAGGTTCCAGA 120
 Db 245 AGTCTTCTTAAATTGCCGGAAGTAGGGAATTCGCCGCACTGCAAGCCCAAAGGTCGAGAGA 304
 Qy 121 TGGGCTAGTGTTAATTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
 Db 305 TGGGCAAGTGTAATTAGGTATCTTTATATGATGCAATGTTCTGGGATCCACAGAGT 364
 Qy 181 CTCGGGGTACACATATCGAAGGTTTCGATCTGCGCACTCTGGACACATGCGTCCCCGAGCG 240
 Db 365 CTTGGGGTACACATATCGAAGGTTTCGATCTGCTACACATAGACACATGCGTTCCTGAGCAG 424
 Qy 241 GTTCGATTTATACAGTCAATCGGAAATCATTAAGCAAAATAGTTACTCGGAAGCAGAGCTA 300
 Db 425 GTTGCTTTTATTCATCAATGGGGAATGAAGAAGCAACAGTTACTCGGAAGCTGAGTTA 484
 Qy 301 CCCCCAAACTATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAGATGATGAAGAGAG 360
 Db 485 CCCCTTAATTATGATAGAGTTGGAAATTCAGAACTTCATCCGTGCAAAAGTATGAGGAAAA 544
 Qy 361 AGATGGGTTTCTAGAGGGGAAAAAGGCTAGATCACTCTCTAGAGTGCAGCAGGAACGG--- 417
 Db 545 AGATGGGTTCTTAGAGATGGAATAATCCAAATCACCACCTAGAAAGGTTGCGAGAAAGGGCT 604
 Qy 418 CGGAAATCTGTGGAGAGAGTGGGCGGGATATGACATGACATGATGATAGTACTGCTCTGTA 477
 Db 605 CCCTCAATTGCGAGAGACCTAATGAACAAGTGGTCAATGGGCACATTATTAAATCCGAG 664
 Qy 478 AATTGTTTGGAGGAGGAAAACTATTCCAGCATCTAGAACAAAGAAAAATGATGTTGCTGCA 537
 Db 665 AATTCAATTGAGGAAGAGGAAACAAACAGCACTTGGTCAGAAAGAAATCTTCTCTGCA 724

AY596550 2173 bp mRNA linear HTC 30-MAY-2004
 Saccharum officinarum clone SCCRZ1001B01, complete sequence.
 DEFINITION
 AY596550
 VERSION
 AY596550.1 GI:47605274
 HTC
 KEYWORDS
 Saccharum officinarum
 ORGANISM
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 REFERENCE
 1 (bases 1 to 2173)
 Oliveira, L.P., Carvalho, R.A., Drezza, A.L., Fregolente, M.C.,
 Martini, I.J., Nascimento, E.O., Tsuneda, S.S., Rodrigues-Filho, P.C.,
 Ulian, E.C., Nogueira, F.T.S., Vicentini, R., Felix, J.M. and
 Menossi, M.
 Characterization of conserved hypothetical proteins from sugarcane
 (Saccharum sp.)
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 2173)
 Oliveira, L.P., Carvalho, R.A., Drezza, A.L., Fregolente, M.C.,
 Martini, I.J., Nascimento, E.O., Tsuneda, S.S., Rodrigues-Filho, P.C.,
 Ulian, E.C., Nogueira, F.T.S., Vicentini, R., Felix, J.M. and
 Menossi, M.
 Direct Submission
 Submitted (12-APR-2004) Departamento de Genetica e Evolucao,
 Instituto de Biologia, Centro de Biologia Molecular e Engenharia
 Genetica, Universidade Estadual de Campinas - Campinas (UNICAMP),
 Campinas, SP 13083-970, Brazil
 JOURNAL
 Location/Qualifiers
 FEATURES
 source
 1. 2173
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCRZ1001B01"
 ORIGIN
 Query Match 23.6%; Score 343.2; DB 3; Length 2173;
 Best Local Similarity 61.2%; Pred. No. 7.7e-89;
 Matches 656; Conservative 0; Mismatches 383; Indels 33; Gaps 5;
 1 ATGACGAGAAAGCAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
 287 ATGACGAGAAAGGCGTCTGTTCCAGAGCTCAACGCCAAGCAAGAGATATTGGA 346
 61 GGGCTTCTTAACATCCAGAGAAACAGAGAATGTCTGACTGCAAAACAAAAGGTCCA 120
 347 AGTCTTCTCGGCTGCTGAGATAGAGATGCGCGACTGCAAGTCAAGGGTCTCTGA 406
 121 TGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGAGT 180
 407 TGGCAAGTGTGAATCTAGGATCTTTATATGATGACATGTTCTGGAATTCATAGA 466
 181 CTCGGGTACACATATCGAAGTTCTGCTGCTGCACTCTGGACATCGCTCCCGAGCAG 240
 467 CTGGGGTGACATATCTAAGGTGAATCTGCCACCTGGATACATCGCTCCGAGCAA 526
 241 GTTCATTATACAGTCAATGGGAAATGATAAAGCAATAGTTACTGGGAAGCAGACTA 300
 527 GTTGCATTTATCAATCAATGGGAAACGAAAGCAAAATAGCTATTGGGAAGCAGACTG 586
 301 CCCCCAACTATGATAGATTGGAAATTTATATAGTGTGCAAAAGTATGAGAGAAG 360
 587 CCTCCTAACTATGATAGATTGGGATAGAAATTTATCGGTGCAAAATACGAGGAAG 646
 361 AGATGGTCTTACAGGGGAAAGGCTAGATCATCTCTAGATCGAGCAGGACGGCGG 420
 647 AGATGGGTACCGAGGAATGGAACATCAGACCTTCTCCGGTGTTCAGATGAGAAAGC 706
 421 AAACTGTGGAGAGAAGTGGCGGGGATATGAGCATGGACATAGTAGT--AGTCTGTAA 478
 707 CAAGAGTCTCGGCTAGTGTCTAATAGGAGTGGACATGTCATAGATCTTCATTGAGCAA 766

479 ATTTGTTTGGAGAGAGAAACTATTTCAGCATCTAGAACAGAAATAATTTGTGCGAA 538
 767 AACCGTGGTTACCAGCTCTTCGAGCAAGTTGCACCTGTAGCTTCAAGGATACCCCTCT 826
 539 CGAGAATAAA-TCTTCCCGTGCCTCCCAAGACCCAGTCAGGTTATAAAGCCACAGCAG 597
 827 CAGGCATCACCTCAGCCCCCAAGGTTAGAACCAACCACTAGTTTCTTAAGGTGGTTTCACTCTCT 886
 598 AAAATGGAGTCTGAGCTACTCCAGTAGAGAGGGAGAAAACAGCAGTAATAATTTGACCA 657
 887 CAGCCACAGAAATCTCTCCAAAGTTGAGGCAACACCCCTTAAGTTGAGAGCCATCT 946
 658 GCATCAGATCTCCAAAGGTGGATTGCTACTGATCTGTTTAAATGCTATCAATGGAT 717
 947 GTTGGCCCACTCTAAAGTTGATTATGCTACTGATCTCTTTAAATGTTTATCAATGGAT 1006
 718 GATTGACTACAAATACCTCAGAGGCACTCTGGGATACTCTCTCCGATGATACTCA 777
 1007 GGAACAAACAGAGAAAGAGTCAGAGTCATCTTCAAAC-----GATGATAATGCC 1054
 778 TGGGCTGGCTTTCAGTCTGCTGNAAGTGTCAAACGCGAGAGAAAATTTGCACAGCAAG 837
 1055 TGGATGGCTTTCAGTCTGCAACCAAGTACCTAGCTCAGAGAAAAGATTCTGCGCAA 1114
 838 CTGCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTTGAGGATTTGTTAAGGAC 897
 1115 CCAGCAGAAAGCAAGACCC-----AGTCAACATCTGGAATAGAGACTTATTATAAGAC 1168
 898 ACACCTAATTTAACTCAACTCAACAGCACC-----AAAAGTGTGAAAGGCGAT 945
 1169 TCACCGCTGTGTCAATATATCTCAGCTCCAGCTGTTTCCCAAGTAATGCGAAGATGAT 1228
 946 ATCATGAGCTGTTTTCAGAGAGAGATATAGTATCGCTTTTGCCATGCATCAGCAACAG 1005
 1229 ATCATGAGTTGTTTGAAGATCCAATATGTTATCTCCATTGCTGCCCATCAGCAGCAG 1288
 1006 GTTGTCTATGCTGCTCAGCAGCAAGCCCTTTATCATGCTGCGAGCAAGAGCTG 1057
 1289 CTGGCGTTATGTTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTTAAAGCTG 1340
 CV197356 742 bp mRNA linear EST 15-SEP-2004
 CGF1003951_F06 Seed coat from mid-season walnut embryos collected
 Aug 1 Juglans regia cDNA clone WSC0009_IF_F06 5', mRNA sequence.
 CV197356
 VERSION
 CV197356.1 GI:52126193
 EST
 SOURCE
 Juglans regia (English walnut)
 ORGANISM
 Juglans regia
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids 1; Fagales; Juglandaceae; Juglans.
 REFERENCE
 1 (bases 1 to 742)
 Muir, R., Baek, J., Leslie, A., Cook, D. and Dandekar, A.
 Analysis of genes expressed in walnut seed coat tissue
 JOURNAL
 Unpublished (2004)
 CONTACT: Abhaya Dandekar, PhD
 CAES Genome Facility
 UC Davis, Department of Pomology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 752 7784
 Fax: 530 752 8502
 Email: amandekar@ucdavis.edu
 Seq primer: WSCF-TCCGAGATCTGACGAGC.
 Location/Qualifiers
 FEATURES
 source
 1. 742
 /organism="Juglans regia"
 /mol_type="mRNA"
 /cultivar="Tulare"
 /db_xref="taxon:51240"

/clone="WSC0009 IF F06"
 /sex="Hermaphrodite"
 /dev Stage="Mid season fruit collected Aug 1"
 /lab host="Xl10-Gold"
 /clone lib="Seed coat from mid-season walnut embryos collected Aug 1"
 /notes="Organ: Seed coat; Vector: pTriplEx2; Site 1: Sf1A; Site 2: Sf1B; Walnut nut samples were harvested from Tulare trees growing in the 'Stuke Block' in the Wolfskill experimental orchard located in Winters, California (USA). Samples were harvested on August 1, 2001 between 8 and 10 am. Samples were then dissected the same day. Seed coat (pellicle) tissue was separated from embryos and frozen immediately in liquid nitrogen and stored at -80C. A gram of sample was removed and ground to a fine powder in liquid nitrogen. Total RNA was extracted using the hot borate procedure. Poly A+ RNA was obtained using the Poly(A) Purist kit (Ambion). The cDNA library was constructed using the SMART cDNA library kit (Clontech). Primary library was en masse evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (Xl10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/tray), picked using a Qbot and archived in 384 well dishes."

ORIGIN

Query Match 23.4%; Score 339.8; DB 7; Length 742;
 Best Local Similarity 69.8%; Pred. No. 5.3e-88;
 Matches 512; Conservative 0; Mismatches 198; Indels 24; Gaps 3;

QY 1 ATCAAGGAGGAGGCGCTCTTAAGGAGCTTAATGCCCGCATAGAGATCTTTGAA 60
 DB 26 ATGAACGAGAGGCGCAACGTTTCCAAGAGCTCAACGCAAGCACAGAAAGATCTGGAA 85
 QY 61 GGGCTCTTTAAACATCAGAGACAGAGATGTGCTGACTGCAAAACAAAAGTCCCAAG 120
 DB 86 GGACTTTTAAATGCCAGAGATAGGAGATGTGCTGACTGCAAACTAAAGTCCCAAG 145
 QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
 DB 146 TGGGCTAGTGTGAATTTAGGCATATTTATATGTATGTCAGTGTTCGGGGATCCATAGAAT 205
 QY 181 CTCGGGGTACACATATCGAAGTTCGATCTGCCACTCTGCACATGCTCCCGAGCAG 240
 DB 206 CTTGGGGTACACATATCGAAGTTCGGGTCTGCTACCTCGACACTTGGCTTCCAGAGCAG 265
 QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTACTCGGAAGCAGAGCTA 300
 DB 266 GTTGCTTTTATTCATTCATGGGAATGAGAGGCAATAGTTACTTGGGAGCCGAACCTA 325
 QY 301 CCCCCAACTATCATAGAGTTGGAATTTATAGTTCAGTCCAAAGTATGAAGAGAG 360
 DB 326 CTTCCAAATTTATGATAGAGTTGGAATTTGAAATTTTATTCGTGCAAAAGTACGATGACAAG 385
 QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGTACCTCTCTAGATCTGAGCAGGAAACGG--- 417
 DB 386 AGATGGATTTCCAAAGGATGGAAAAACAAATATCACTAGAGGGCAGGAAGCAAGGCT 445
 QY 418 CGGAATCTGTGGAGAGAGTGGCGGGGATATGAGCATGAGCATAGTATAGTATCTCTGTA 477
 DB 446 TCTGTGATTTGGCAGAGACTGGGGGAAGAGAGTGGTCCGGGCCCTAGTTCTGTTCTGAA 505
 QY 478 AATTTGTTTGGAGAGGAGAACTATTTCAGCATCTTAGAACAGAGAAATATGTTGTGCA 537
 DB 506 AAGACATTTTGGAGAAAGGAGAACACTCGACCATCTAGTACTA-----TTCCCTCT 556
 QY 538 ACCAGATATAATCTTCCCGTCTCCCAAGGAGCCAGTCAGGTTATAAGCCACAGCAG 597
 DB 557 ACCAAATAGGACATCTCTGTTCACCTTAAAGGACCTGAGCGAGAAATAGAACCCAGCAC-- 614
 QY 598 AAAATGGAGTCTGCAGCTACTCCAGTAGAGGGGAGAAACNAGCAGTAAATGTTGCAACCA 657

DB 615 -----CAGCGTGCCTCCCAAGCTGAAGCAACAAGCAGCCTTTGAATGCTGCTCT 664
 QY 658 GCATCAGATCTCCAAAGTGGATTTTGTCTACTGTCTGATCTGTTTAAACATGCTATCAATGGAT 717
 DB 665 GCTGTCTCTCCCAAGGTTGACTGANGCAACTGACCTTTTCGATATGTTAACCATGGAT 724
 QY 718 GATTGACTACAAA 731
 DB 725 GCTCCAAGTGAAA 738
 RESULT 15
 AY105375
 LOCUS PC0146209 mRNA linear HTC 16-OCT-2002
 DEFINITION Zea mays
 ACCESSION AY105375
 VERSION AY105375.1 GI:21208453
 KEYWORDS HTC.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 2189)
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.
 TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 2189)
 AUTHORS Coe, E.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maize-map.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.
 FEATURES
 Location/Qualifiers
 1..2189
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:638642"
 /clone_lib="Maize Mapping Project/DuPont Consensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 23.3%; Score 338.4; DB 3; Length 2189;
 Best Local Similarity 60.9%; Pred. No. 2e-87;
 Matches 653; Conservative 0; Mismatches 386; Indels 33; Gaps 5;

QY 1 ATGACGAGAGAGGCAAGCTCTCTAAGGAGCTTAATGCCCGCATAGAGATCTTTGAA 60
 DB 299 ATGAACGAGAAAGGCGTCCGTCTCCAGGAGCTCAACGCCACGACAGAGATATTGGAA 358
 QY 61 GGGCTCTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTCCCAAG 420
 DB 359 GGTCTTCTACGCATCTCTGAGAATAGAGATGCGCAGACTGCAAGTCAAAAGGCTCTCGA 418
 QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
 DB 419 TGGGCAAGTGTGAATCTAGGTATCTTTATGATGATGATGTTCTGGCATTATAGAAGC 478

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:39:09 ; Search time 779 Seconds
(without alignments)
11033.972 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaagccacagt.....gaatgttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a
-score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1 | 1452 | 100.0 | 1452 | 6 | AAL48968 | Aal48968 A thalian |
| 2 | 1452 | 100.0 | 1452 | 12 | ADN72208 | Adn72208 Thale cre |
| 3 | 1452 | 100.0 | 1791 | 3 | AAC48822 | Aac48822 Arabidops |
| 4 | 1450.4 | 99.9 | 1452 | 6 | AAL48969 | Aal48969 A thalian |
| 5 | 1450.4 | 99.9 | 1452 | 6 | AAL48970 | Aal48970 A thalian |
| 6 | 1447.2 | 99.7 | 1794 | 3 | AAC38070 | Aac38070 Arabidops |
| 7 | 298.2 | 20.5 | 596 | 13 | ACN48712 | Acn48712 Cotton pr |
| 8 | 255.4 | 17.6 | 5580 | 6 | AAL48972 | Aal48972 A thalian |
| 9 | 237.2 | 16.3 | 371 | 3 | AAC75657 | Aac75657 Human ORF |
| 10 | 237.2 | 16.3 | 371 | 6 | ABN26360 | Abn26360 Human ORF |
| 11 | 227.4 | 15.7 | 313 | 6 | ABL71943 | Abl71943 Corn tass |
| 12 | 199.6 | 13.7 | 270 | 10 | ABX82114 | Abx82114 Corn ear- |
| 13 | 182.4 | 12.6 | 408 | 3 | AAC56661 | Aac56661 Eucalyptu |
| 14 | 127.2 | 8.8 | 290 | 6 | ABL73349 | Abl73349 Corn tass |
| 15 | 107.2 | 7.4 | 587 | 6 | ABQ98937 | Abq98937 Human ORF |
| 16 | 107.2 | 7.4 | 3117 | 4 | AAC84451 | Aac84451 Nucleotid |
| 17 | 105.6 | 7.3 | 799 | 4 | AAN06694 | Aan06694 Human cDN |
| 18 | 105.6 | 7.3 | 2314 | 4 | AAH17887 | Aah17887 Human cDN |
| 19 | 105.6 | 7.3 | 2314 | 8 | ACD13426 | Acd13426 Human DNA |
| 20 | 98.4 | 6.8 | 1403 | 12 | ADJ57942 | Adj57942 Rat NARC |

| | | | | | | |
|----|------|-----|------|----|-----------|--------------------|
| 21 | 94.8 | 6.5 | 2561 | 4 | AAC84446 | Aac84446 Nucleotid |
| 22 | 94.2 | 6.5 | 667 | 4 | AAI60996 | Aai60996 Human pol |
| 23 | 94.2 | 6.5 | 1447 | 6 | ABS70464 | Abs70464 Human bon |
| 24 | 94.2 | 6.5 | 1976 | 4 | AAI58627 | Aai58627 Human pol |
| 25 | 94.2 | 6.5 | 1976 | 5 | ADQ98845 | Adq98845 DNA encod |
| 26 | 94.2 | 6.5 | 1976 | 9 | ADB48605 | Adb48605 Novel hum |
| 27 | 94.2 | 6.5 | 2457 | 8 | ACD13427 | Acd13427 Human DNA |
| 28 | 94.2 | 6.5 | 2492 | 3 | AACT76483 | Aac76483 Human ORF |
| 29 | 94.2 | 6.5 | 2618 | 13 | ADP55646 | Adp55646 Human PRO |
| 30 | 94.2 | 6.5 | 2874 | 12 | ADQ97937 | Adq97937 Human can |
| 31 | 92.2 | 6.3 | 2541 | 2 | AAO4343 | Aao4343 Human sec |
| 32 | 86.2 | 5.9 | 309 | 3 | AAO2721 | Aao2721 Human sec |
| 33 | 82.6 | 5.7 | 1327 | 4 | AAI59210 | Aai59210 Human pol |
| 34 | 82.6 | 5.7 | 1327 | 5 | ADQ99434 | Adq99434 DNA encod |
| 35 | 82.6 | 5.7 | 1327 | 9 | ADB49194 | Adb49194 Novel hum |
| 36 | 82.6 | 5.7 | 2176 | 4 | AAI60413 | Aai60413 Human pol |
| 37 | 82 | 5.6 | 1324 | 12 | ADP22613 | Adp22613 Sea-squir |
| 38 | 82 | 5.6 | 1356 | 12 | ADQ08635 | Adq08635 Ciona int |
| 39 | 80.6 | 5.6 | 1274 | 3 | AAC39358 | Aac39358 Arabidops |
| 40 | 80.4 | 5.5 | 507 | 10 | ADG37966 | Adg37966 Aspergill |
| 41 | 80.4 | 5.5 | 1314 | 3 | AAC66455 | Aac66455 Human sec |
| 42 | 80.4 | 5.5 | 2545 | 4 | AAS02552 | Aas02552 Human sec |
| 43 | 79.4 | 5.5 | 555 | 13 | ACN49796 | Acn49796 Cotton pr |
| 44 | 79.2 | 5.5 | 1775 | 12 | ADP22611 | Adp22611 Sea-squir |
| 45 | 78.6 | 5.4 | 398 | 9 | ACH29095 | Ach29095 Human adu |

ALIGNMENTS

RESULT 1
AAL48968
ID AAL48968 standard; DNA; 1452 BP.
XX
AC AAL48968;
XX
DT 28-OCT-2002 (first entry)
XX
DE A thaliana NEVERSHED protein coding sequence.

XX NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
KW GTP-binding protein domain; organ abscission; floral abscission; plant;
KW gene; ds.
XX Arabidopsis thaliana.
XX OS

XX FH Key Location/Qualifiers
XX CDS 1..1452
XX FT /*tag= a
XX FT /product= "NEVERSHED"
XX
XX PN WO200261042-A2.

XX PD 08-AUG-2002.
XX
XX PF 22-JAN-2002; 2002WO-US001938.
XX
XX PR 29-JAN-2001; 2001US-0264974P.
XX
XX PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX (REGC) UNIV CALIFORNIA.
XX
XX PI Liljegren SJ, Becker JR, Yanofsky MF;
XX WPI; 2002-627475/67.
XX P-PSDB; AAO18759.

XX PT New mutant plants comprising a modified ARF-GAP domain and having
XX decreased organ abscission, particularly floral abscission.
XX Claim 11; Page 44-46; 67pp; English.
XX The present invention relates to a mutant plant which has decreased organ
CC

CC abscission, comprising a nucleotide sequence having a modified ARF GAP
 CC domain. The mutant plant is useful as a source of plants or progenies
 CC having reduced or decreased floral abscission. The present sequence is
 CC the Arabidopsis thaliana NEVERSHED protein coding sequence, which can be
 CC mutated in plants of the invention
 XX

SQ Sequence 1452 BP; 461 A; 333 C; 329 G; 329 T; 0 U; 0 Other;

Query Match 100.0%; Score 1452; DB 6; Length 1452;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAGAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATCTCTGAA 60
 DB 1 ATGAACGAGAGAGAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATCTCTGAA 60
 QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGACTGTGCAAAACAAAGAGTCCAAGA 120
 DB 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGACTGTGCAAAACAAAGAGTCCAAGA 120
 QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
 DB 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
 QY 181 CTCGGGTTACATATCGAAGGTTTCGATCTGCACTCTGGACACATGGCTCCCGCAGCAG 240
 DB 181 CTCGGGTTACATATCGAAGGTTTCGATCTGCACTCTGGACACATGGCTCCCGCAGCAG 240
 QY 241 GTTGCAATTATACATCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
 DB 241 GTTGCAATTATACATCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
 QY 301 CCCCAAACTATAGTAGTTCGAAATGAGATTTTATACGTGCAAGTATGAAAGAG 360
 DB 301 CCCCAAACTATAGTAGTTCGAAATGAGATTTTATACGTGCAAGTATGAAAGAG 360
 QY 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCGG 420
 DB 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTCTAGAGTCGAGCAGGAAACGGCGG 420
 QY 421 AAATCTGTGGAGAGAGTGGCGGGGATATAGCATATGAGCATATGATAGTAGTCTCTAAAT 480
 DB 421 AAATCTGTGGAGAGAGTGGCGGGGATATAGCATATGAGCATATGATAGTAGTCTCTAAAT 480
 QY 481 TTGTTTGGAGAGAGGAAATCTATCCAGCATCTAGAAACAAAGAAATAGTTGCTGCAACG 540
 DB 481 TTGTTTGGAGAGAGGAAATCTATCCAGCATCTAGAAACAAAGAAATAGTTGCTGCAACG 540
 QY 541 AGAATAAATCTTCCCGTGCCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
 DB 541 AGAATAAATCTTCCCGTGCCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
 QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAATTTGTCACAGCA 660
 DB 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAATTTGTCACAGCA 660
 QY 661 TCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
 DB 661 TCAGATCTCCAAAGGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
 QY 721 TCGACTACAAATACCTTCAGAGGCAACTCCTGGCGATCTCTCCCGGATGATTAATCATGG 780
 DB 721 TCGACTACAAATACCTTCAGAGGCAACTCCTGGCGATCTCTCCCGGATGATTAATCATGG 780
 QY 781 GCTGGCTTTTCACTGTCTGGAAGTGTGTTCAAAACGGCAGAGAAAATTTGTCAAGCCAGCT 840
 DB 781 GCTGGCTTTTCACTGTCTGGAAGTGTGTTCAAAACGGCAGAGAAAATTTGTCAAGCCAGCT 840
 QY 841 GCTGAGAGAGGTTCTCTCCAGCTTCATCTTCTGACTTTGAGATTTGTTTAAAGACACA 900
 DB 841 GCTGAGAGAGGTTCTCTCCAGCTTCATCTTCTGACTTTGAGATTTGTTTAAAGACACA 900
 QY 901 CCTAATTTTAACTCAACAGCACCACAAAGATGTGAAAGGCGATATCATGAGCCTGTTT 960

DB 901 CCTAATTTTAACTCAACAGCACCACAAAGATGTGAAAGCGGATATCATGAGCCTGTTT 960
 QY 961 GAGAAGAGAGATATAGTATGCGCTTTTGGCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
 DB 961 GAGAAGAGAGATATAGTATGCGCTTTTGGCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
 QY 1021 CAGCAGCAAGCCCTTTACATGGCTGCGAGCGAAAGCTGCTGAGGCACTCCAAAACGGCGTG 1080
 DB 1021 CAGCAGCAAGCCCTTTACATGGCTGCGAGCGAAAGCTGCTGAGGCACTCCAAAACGGCGTG 1080
 QY 1081 AATCAACAAGCTATTTGCTTAACTGCTTTTAAAGTCTTCTGCAAAATTTGTCAAAACCCCGGC 1140
 DB 1081 AATCAACAAGCTATTTGCTTAACTGCTTTTAAAGTCTTCTGCAAAATTTGTCAAAACCCCGGC 1140
 QY 1141 GGCTACAGATCCCGGAATGACTTAAACCCGCTAGGTGCTCAAGCTGATCTCCAGAACTT 1200
 DB 1141 GGCTACAGATCCCGGAATGACTTAAACCCGCTAGGTGCTCAAGCTGATCTCCAGAACTT 1200
 QY 1201 ATGCAAAACATGAATATGAACGCAAAACATGAACACGAGACCCGCGACAAACCGCAAGAGAAC 1260
 DB 1201 ATGCAAAACATGAATATGAACGCAAAACATGAACACGAGACCCGCGACAAACCGCAAGAGAAC 1260
 QY 1261 ACTCTAACAATACCATCATCCAGTTTCTACAAATGGGTCAAGCTATCAAGTGAACGGT 1320
 DB 1261 ACTCTAACAATACCATCATCCAGTTTCTACAAATGGGTCAAGCTATCAAGTGAACGGT 1320
 QY 1321 ATGACCCCAACTCAACCCGCTTAAACCTCAGTCATCATCCGCAACCCCAACCAAGCACC 1380
 DB 1321 ATGACCCCAACTCAACCCGCTTAAACCTCAGTCATCATCCGCAACCCCAACCAAGCACC 1380
 QY 1381 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTTCTTGTGATGGAATGTTTC 1440
 DB 1381 ACACCATCTTCACAATCAGGCAAAAGACTTTGATTTCTTCTTGTGATGGAATGTTTC 1440
 QY 1441 ACAAAACATTTGA 1452
 DB 1441 ACAAAACATTTGA 1452

RESULT 2

ADN72208
 ID ADN72208 standard; cDNA; 1452 BP.

XX ADN72208;

XX 15-JUL-2004 (first entry)

XX Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 103.

XX gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
 XX growth regulator; animal feed product; thale cress;
 XX cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

XX Inze D, De Veyder L, Vlieghe K;

XX WPI; 2004-348466/32.

XX P-PSDB; ADN72209.

XX Altering plant characteristics, useful for producing plants for enzyme or
 PT pharmaceutical production comprises modifying in a plant, expression of

one or more nucleic acids and/or modifying level or activity of one or more proteins.

Claim 1: SEQ ID NO 103; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, signal architecture or physiology, altered endoreduplication, biochemistry, plant transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress CDNA upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa transcription factor, given in an exemplification of the invention.

Query Match 100.0%; Score 1452; DB 12; Length 1452;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | ATGACGAGAAAGCCAACTCTTAAGGAGCTTAATGCGCGCCATAGAAAGATTCTTGAA | 60 |
| DB | 1 | ATGACGAGAAAGCCAACTCTTAAGGAGCTTAATGCGCGCCATAGAAAGATTCTTGAA | 60 |
| QY | 61 | GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAGAGTCCAAGA | 120 |
| DB | 61 | GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAGAGTCCAAGA | 120 |
| QY | 121 | TGGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT | 180 |
| DB | 121 | TGGGCTAGTGTAAATTTAGGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT | 180 |
| QY | 181 | CTCGGGGTACACATATCGAAGGTTCCATCTGCCACTCGGACATGCGCTCCCGAGCAG | 240 |
| DB | 181 | CTCGGGGTACACATATCGAAGGTTCCATCTGCCACTCGGACATGCGCTCCCGAGCAG | 240 |
| QY | 241 | GTTCGATTTTATACAGTCAATGGGAAATGATAAGCAATAGTTACTTGGGAAGCAGAGCTA | 300 |
| DB | 241 | GTTCGATTTTATACAGTCAATGGGAAATGATAAGCAATAGTTACTTGGGAAGCAGAGCTA | 300 |
| QY | 301 | CCCCAACTATGATAGAGTTGGAATTGAGAATTTTATACGTGCAAGATGATGAAGAGAG | 360 |
| DB | 301 | CCCCAACTATGATAGAGTTGGAATTGAGAATTTTATACGTGCAAGATGATGAAGAGAG | 360 |
| QY | 361 | AGATGGGTTCTAGAGGGGAAAGGCTAGTACATCTCTAGAGTCAGCAGGAGAACGGGG | 420 |
| DB | 361 | AGATGGGTTCTAGAGGGGAAAGGCTAGTACATCTCTAGAGTCAGCAGGAGAACGGGG | 420 |
| QY | 421 | AAATCTGTGAGAGAGTGGCGGGATATGACATGGGACATAGTAGTCTGTAAAT | 480 |
| DB | 421 | AAATCTGTGAGAGAGTGGCGGGATATGACATGGGACATAGTAGTCTGTAAAT | 480 |
| QY | 481 | TTGTTTGGAGAGGAAAACTATTCAGCATCTTAGAACAGAAATATGTTGCTGCAACG | 540 |
| DB | 481 | TTGTTTGGAGAGGAAAACTATTCAGCATCTTAGAACAGAAATATGTTGCTGCAACG | 540 |
| QY | 541 | AGAAATAATCTTCCCGTGGCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGAGAA | 600 |
| DB | 541 | AGAAATAATCTTCCCGTGGCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGAGAA | 600 |
| QY | 601 | ATGGAGCTGCGAGTACTCTCAGTAGAGGGGAGAAACAAAGCAGTAATGTTGCACAGCA | 660 |
| DB | 601 | ATGGAGCTGCGAGTACTCTCAGTAGAGGGGAGAAACAAAGCAGTAATGTTGCACAGCA | 660 |

| | | | |
|----|------|---|------|
| DB | 601 | ATGGAGCTGCGAGTACTCTCAGTAGAGGGGAGAAACAAAGCAGTAATGTTGCACAGCA | 660 |
| QY | 661 | TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATCTATCAATGATGAT | 720 |
| DB | 661 | TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATCTATCAATGATGAT | 720 |
| QY | 721 | TCGACTACAAATACCTCAGAGGCAATCTCTGGGATATCTCTGCGATGATCAATCATGG | 780 |
| DB | 721 | TCGACTACAAATACCTCAGAGGCAATCTCTGGGATATCTCTGCGATGATCAATCATGG | 780 |
| QY | 781 | GCTGGCTTTCAGTCTGCTGGAGTGGTCAACGGCAGAGAAAATTTGTCAAGCCAAAGCT | 840 |
| DB | 781 | GCTGGCTTTCAGTCTGCTGGAGTGGTCAACGGCAGAGAAAATTTGTCAAGCCAAAGCT | 840 |
| QY | 841 | GCTGAGAGCAGTCTCTCTCCAGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGACACA | 900 |
| DB | 841 | GCTGAGAGCAGTCTCTCTCCAGCTTTCATCTTCTGACTTTGAGGATTTGTTTAAAGACACA | 900 |
| QY | 901 | CCTAATTTTAACTCAACAGGCAACCAAGATGTGAAGGCGATATCATGAGCCTGTTT | 960 |
| DB | 901 | CCTAATTTTAACTCAACAGGCAACCAAGATGTGAAGGCGATATCATGAGCCTGTTT | 960 |
| QY | 961 | GAGAGCAGATATAGTATGCTTTCCTTTCCTCATGATCAGCAAGGTTGCTATGCTCGCT | 1020 |
| DB | 961 | GAGAGCAGATATAGTATGCTTTCCTTTCCTCATGATCAGCAAGGTTGCTATGCTCGCT | 1020 |
| QY | 1021 | CAGCAGCAAGCCCTTTTACATGGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG | 1080 |
| DB | 1021 | CAGCAGCAAGCCCTTTTACATGGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG | 1080 |
| QY | 1081 | AATCAACAAGCTATTTGCTTAATGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTT | 1140 |
| DB | 1081 | AATCAACAAGCTATTTGCTTAATGCTTAACTGCTTAACTGCTTAACTGCTTAACTGCTT | 1140 |
| QY | 1141 | GGCTACGAGATCCCGGATGACTAAACCCGCTAGTGTCTCAAGCTGATCTCCAGAACTT | 1200 |
| DB | 1141 | GGCTACGAGATCCCGGATGACTAAACCCGCTAGTGTCTCAAGCTGATCTCCAGAACTT | 1200 |
| QY | 1201 | ATGCAGAAACATGATATGAAACGCAACATGAAACGAGACCCGCAACCCGCAAGAGAAC | 1260 |
| DB | 1201 | ATGCAGAAACATGATATGAAACGCAACATGAAACGAGACCCGCAACCCGCAAGAGAAC | 1260 |
| QY | 1261 | ACTCTACATATCCCATCATCTCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGCT | 1320 |
| DB | 1261 | ACTCTACATATCCCATCATCTCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGCT | 1320 |
| QY | 1321 | ATGACCCCAAACTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAACAGCACC | 1380 |
| DB | 1321 | ATGACCCCAAACTCAACCGGTAAACCTCAGTCAATCCGCAACCCCAACCAACAGCACC | 1380 |
| QY | 1381 | ACACCATCTTCAATCAGGCAAGACCTTGAATTTCTCTTCTTGTGATGGAATGTTTC | 1440 |
| DB | 1381 | ACACCATCTTCAATCAGGCAAGACCTTGAATTTCTCTTCTTGTGATGGAATGTTTC | 1440 |
| QY | 1441 | ACAAACATTTGA | 1452 |
| DB | 1441 | ACAAACATTTGA | 1452 |

RESULT 3

AAC48822

ID AAC48822 standard; DNA; 1791 BP.

XX AAC48822;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58899.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

XX promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD
XX PF 06-SEP-2000.
XX PP 25-FEB-2000; 2000BP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 18-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132487P.
PR 14-MAY-1999; 99US-0132663P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

[illegible]

RESULT 4
 AAL48969 ID AAL48969 standard; DNA; 1452 BP.
 XX AC AAL48969;
 XX DT 28-OCT-2002 (first entry)
 XX DE A thaliana EMS mutated NEVERSHED protein coding sequence mutant nev-1.
 XX KW NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
 XX KM GTP-binding protein domain; organ abscission; floral abscission; plant;
 XX KW mutant; nev-1; gene; ds.
 XX OS Arabidopsis thaliana.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX CDS 1..1452
 XX FT /*tag= a
 XX FT /product= "nev-1"
 XX PN W0200261042-A2.
 XX PD 08-AUG-2002.
 XX PF 22-JAN-2002; 2002W0-US001938.
 XX PR 29-JAN-2001; 2001US-0264974P.
 XX (SALK) SALK INST BIOLOGICAL STUDIES.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PI Liljegen SU, Ecker JR, Yanofsky MF;
 XX WPI; 2002-627475/67.
 XX P-PSDB; AAO18760.
 XX PT New mutant plants comprising a modified ARF-GAP domain and having
 XX PT decreased organ abscission, particularly floral abscission.
 XX PS Claim 6; Page 48-50; 67pp; English.
 XX CC The present invention relates to a mutant plant which has decreased organ
 XX CC abscission, comprising a nucleotide sequence having a modified ARF GAP
 XX CC domain. The mutant plant is useful as a source of plants or progenies
 XX CC having reduced or decreased floral abscission. The present sequence is
 XX CC the Arabidopsis thaliana NEVERSHED protein mutant coding sequence nev-1,
 XX CC which can be used in plants of the invention
 XX SQ Sequence 1452 BP; 462 A; 333 C; 328 G; 329 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1450.4; DB 6; Length 1452;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAACGAGAGCCAAACGCTCTTAAGAGCTTAATGCCCGGCATAGAAAGATTCTTGAA 60
 DB 1 ATGAACGAGAGCCAAACGCTCTTAAGAGCTTAATGCCCGGCATAGAAAGATTCTTGAA 60
 QY 61 GGGCTTTTAAACATCCAGAGACAGAGATGTGCTGACTGCNAAACAAAGGTCCAAGA 120
 DB 61 GGGCTTTTAAACATCCAGAGACAGAGATGTGCTGACTGCNAAACAAAGGTCCAAGA 120
 QY 121 TGGGCTAGTGTAAATTAGTATCTTTATCTGCTCAATGTCTCGGATTCAAGAGT 180
 DB 121 TGGGCTAGTGTAAATTAGTATCTTTATCTACATGCAATGTCTCGGATTCAAGAGT 180
 QY 181 CTCGGGGTACACATATCGAAGGTTCGATCTGCCACTCTCGACACATGGCTCCCGAGCAG 240
 DB 181 CTCGGGGTACACATATCGAAGGTTCGATCTGCCACTCTCGACACATGGCTCCCGAGCAG 240

QY 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 GTTGCAATTTATACAGTCAATGGGAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
 QY 301 CCCCCAAAATATGATAGAGTTGGAAATTGAGAAATTTTATACGTGCAAAAGTATGAAGAGAAG 360
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 CCCCCAAAATATGATAGAGTTGGAAATTGAGAAATTTTATACGTGCAAAAGTATGAAGAGAAG 360
 QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGATCGATGAGCAGAAACGGCGG 420
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGATCGATGAGCAGAAACGGCGG 420
 QY 421 AAATCTGTGGAGAGAAAGTGGGCGGGGATATGAGCATGACATAGTAGTAGTCTCTTAAT 480
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 421 AAATCTGTGGAGAGAAAGTGGGCGGGGATATGAGCATGACATAGTAGTAGTCTCTTAAT 480
 QY 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGACAAGAAATAATGTTCTCTGCAACG 540
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGACAAGAAATAATGTTCTCTGCAACG 540
 QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAAGTTTATAAGCCACAGCAGAA 600
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAAGTTTATAAGCCACAGCAGAA 600
 QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAAACAGCAGTAATGTTGCACACGCA 660
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAAACAGCAGTAATGTTGCACACGCA 660
 QY 661 TCAGATCTCCAAAGGTGGATTTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 661 TCAGATCTCCAAAGGTGGATTTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
 QY 721 TCGACTACAAATACCTCAGAGGCACTCTCGCGATACCTCTGCGGATGATCAATCATG 780
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 721 TCGACTACAAATACCTCAGAGGCACTCTCGCGATACCTCTGCGGATGATCAATCATG 780
 QY 781 GCTGCTTTTCACTGCTGGAAGTGGTCAAAAGGAGGAGAAAATGTCACGCCAAGCT 840
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 781 GCTGCTTTTCACTGCTGGAAGTGGTCAAAAGGAGGAGAAAATGTCACGCCAAGCT 840
 QY 841 GCTGAGAGAGTTCCTCCAGCTTCATCTTCTGATTTGAGGATTTGTTAAGACACA 900
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 841 GCTGAGAGAGTTCCTCCAGCTTCATCTTCTGATTTGAGGATTTGTTAAGACACA 900
 QY 901 CCTAATTTAAACACTCAACAGCACCAAGATGTGAAGGCGATATCATGAGCCCTGTT 960
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 901 CCTAATTTAAACACTCAACAGCACCAAGATGTGAAGGCGATATCATGAGCCCTGTT 960
 QY 961 GAGAGAGCAATATAGTATCGCTTTTGGCCATGCAATCAGCAACAGGTTGCTATGCTCGCT 1020
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 961 GAGAGAGCAATATAGTATCGCTTTTGGCCATGCAATCAGCAACAGGTTGCTATGCTCGCT 1020
 QY 1021 CAGCAGAAAGCCCTTTAATGCTGTCAGGAAAGCTGCTGGAGGCACTCCAAACGGCGT 1080
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1021 CAGCAGAAAGCCCTTTAATGCTGTCAGGAAAGCTGCTGGAGGCACTCCAAACGGCGT 1080
 QY 1081 AATCAACAGCTATTGCTTAATGCTCTTAAAGTAGCTTCTGCAAAATTTGGTCAAAACCCCGG 1140
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1081 AATCAACAGCTATTGCTTAATGCTCTTAAAGTAGCTTCTGCAAAATTTGGTCAAAACCCCGG 1140
 QY 1141 GGCTACCAAGATCCCCGGGAATGACTAAACCCGCTAGGTGGTCAAGCTGATCTCAGAAACTT 1200
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1141 GGCTACCAAGATCCCCGGGAATGACTAAACCCGCTAGGTGGTCAAGCTGATCTCAGAAACTT 1200
 QY 1201 ATGCAAAAATGAAATGAAACGCAAAATGAAACAGCAGACCCGCAACCCGCAAGAGAAC 1260
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1201 ATGCAAAAATGAAATGAAACGCAAAATGAAACAGCAGACCCGCAACCCGCAAGAGAAC 1260
 QY 1261 ACTCTACAAATCCCATCATCCAGTTTCTACAAATGGGTCAGGCTAATCAAGTGAACGCT 1320
 DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1261 ACTCTACAAATCCCATCATCCAGTTTCTACAAATGGGTCAGGCTAATCAAGTGAACGCT 1320

QY 1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCGCAACCAACCAACGACCC 1380
 DB |||||
 QY 1381 ACACCATCTTCAAAATCAGGCAAGAATTGATTTCTCTCTTGATGATGGAATGTTTC 1440
 DB |||||
 QY 1441 ACAAAACATTGA 1452
 DB |||||
 QY 1441 ACAAAACATTGA 1452
 DB |||||
 RESULT 5
 ID AAL48970
 AC AAL48970 standard; DNA; 1452 BP.
 AC AAL48970;
 DT 28-OCT-2002 (first entry)
 DE A thaliana EMS mutated NEVERSHED protein coding sequence mutant nev-2.
 DE NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
 KW GTP-binding protein domain; organ abscission; floral abscission; plant;
 KW mutant; nev-2; gene; ds.
 KW
 OS Arabidopsis thaliana.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 1..594
 FT /*tag= a
 FT /*product= "nev-2"
 PN WO200261042-A2.
 PD 08-AUG-2002.
 PF 22-JAN-2002; 2002WO-US001938.
 PR 29-JAN-2001; 2001US-0264974P.
 PR (SALK) SALK INST BIOLOGICAL STUDIES.
 PA (REGC) UNIV CALIFORNIA.
 PI Liljgren SJ, Ecker JR, Yanofsky MF;
 DR WPI: 2002-627475/67.
 DR P-PSDB; AAO18761.
 PT New mutant plants comprising a modified ARF-GAP domain and having
 PT decreased organ abscission, particularly floral abscission.
 XX Claim 6; Page 52-53; 67pp; English.
 CC The present invention relates to a mutant plant which has decreased organ
 CC abscission, comprising a nucleotide sequence having a modified ARF GAP
 CC domain. The mutant plant is useful as a source of plants or progenies
 CC having reduced or decreased floral abscission. The present sequence is
 CC the Arabidopsis thaliana NEVERSHED protein mutant coding sequence nev-2,
 CC which can be used in plants of the invention
 XX
 SQ Sequence 1452 BP; 461 A; 332 C; 329 G; 330 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1450.4; DB 6; Length 1452;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGACGAGAGAACCAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 60
 DB |||||
 1 ATGACGAGAGAACCAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 60

QY 61 GGGCTTTCTTAAATCCAGAGAACAGAGAAATGTCTGATCGCAAAACAAAAGTCCCAAGA 120
 DB |||||
 QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
 DB |||||
 QY 181 CTGGGGGTACACATATCGAAGGTTTCACTCTGGACACATGGCTCCCGAGCAG 240
 DB |||||
 QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
 DB |||||
 QY 301 CCCCCAAACTATGATAGAGTTGGAAATTCAGAAATTTTATACGTGCAAGTATCAAGAGAG 360
 DB |||||
 QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTCGAGCAGGAACGGCGG 420
 DB |||||
 QY 421 AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGGAACATAGTAGTAGTCTCTGTAAT 480
 DB |||||
 QY 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGAACAAAGAAATAATGTTTCTGCAACG 540
 DB |||||
 QY 541 AGATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCAGGTTTATAAGCCACAGCAGAAA 600
 DB |||||
 QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAAAGCAGTAAATGTTGCACAGCA 660
 DB |||||
 QY 661 TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
 DB |||||
 QY 721 TCGACTACAAATACCTCAGAGGCACTCTCGCGGATCTCTCGCGGATGATAACTCATGG 780
 DB |||||
 QY 781 GCTGGCTTTTCAGTCTGTGGAAGTGTCAAACGGCAGAGAAAATTTGTCAAGCCAGCCT 840
 DB |||||
 QY 841 GCTGAGAGAGTTCTCTCCAGCTTCACTTCTGATCTTGGAGGATTTGTTTAAAGCACA 900
 DB |||||
 QY 901 CCTAATTTAAACACTCAACAGCACAAAGATGTGAAGGCGATATCAATGAGCGTGT 960
 DB |||||
 QY 961 GAGAAAGAGAAATATAGTATCGCTTTTCCATGCAATCAGCAACAGGTTGCTATGCTGCT 1020
 DB |||||
 QY 1021 CAGCAGAGAGCGCTTTTACATGCTGCGGAAAGCTGCTGGAGGCACTCCAAACGGCGTG 1080
 DB |||||
 QY 1081 AATCAACAGCTATTGCTTAATGCTCTTAAACGTAAGTCTCTGCAAAATTTGGTCAAAACCCCGC 1140
 DB |||||
 QY 1141 GGCTACCAGATCCCGGAAATGATAACCCCGTAGGTGGTCAAGCTGATCTCCAGAAACTT 1200

Db 1141 |||||GGCTACCCAGATCCCGGAATGACTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAAACTT 1200
Qy 1201 ATGCAAAACATGAATATGACCGCAACATGAACGAGACCGGCACACCGCAAGGAAC 1260
Db 1201 ATGCAAAACATGAATATGACCGCAACATGAACGAGACCGGCACACCGCAAGGAAC 1260
Qy 1261 ACTCTACAATACCCATCATCCAGTTTCTTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Db 1261 ACTCTACAATACCCATCATCCAGTTTCTTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Qy 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTATCATTCGCGAACCCCAACCAAGCAACC 1380
Db 1321 ATGACCCCAACTCAACCGGTAAACCTCAGTATCATTCGCGAACCCCAACCAAGCAACC 1380
Qy 1381 ACACCATCTTCAATCAGCGCAAGACTTTTGATTTCTCTTCTGATGATGGAATGTC 1440
Db 1381 ACACCATCTTCAATCAGCGCAAGACTTTTGATTTCTCTTCTGATGATGGAATGTC 1440
Qy 1441 ACAAAACATTGA 1452
Db 1441 ACAAAACATTGA 1452

RESULT 6
AAC38070
ID AAC38070 standard; DNA; 1794 BP.
AC AAC38070;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19684.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137528P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

| | | |
|---|--------------|--|
| PR | 26-OCT-1999; | 99US-0161359P. |
| PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 26-OCT-1999; | 99US-0161361P. |
| PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 28-OCT-1999; | 99US-0161922P. |
| PR | 28-OCT-1999; | 99US-0161993P. |
| PR | 29-OCT-1999; | 99US-0162142P. |
| Query Match 99.7%; Score 1447.2; DB 3; Length 1794; | | |
| Best Local Similarity 99.8%; Pred. No. 0; | | |
| Matches 1449; Conservative 0; Mismatches 3; Indels 0; Gaps 0; | | |
| Qy | 1 | ATGAACGAGAAAGCCCAACGGTCTCTAAGGAGCTTAATCGGCCCATAGAAAGATTCTTTGAA 60 |
| Dd | 167 | ATGAACGAGAAGCCCATGTCTCTAAGGAGCTTAATCGGCCCATAGAAAGATTCTTTGAA 226 |
| Qy | 61 | GCGCTTCTTAAACATCCAGAGAACAGAGATGTGCTACTGCATAAACAAGGTCACAGA 120 |
| Dd | 227 | GCGCTTCTTAAACATCCAGAGAACAGAGATGTGCTACTGCATAAACAAGGTCACAGA 286 |
| Qy | 121 | TGGGCTAGTGTTAATTTAGGTATCTTTTATCTGCATGCAATGTTCTGGATTACAGAGT 180 |
| Dd | 287 | TGGGCTAGTGTTAATTTAGGTATCTTTTATCTGCATGCAATGTTCTGGATTACAGAGT 346 |
| Qy | 181 | CTCGGGGTACACATATCGAAGTTTCGATCTGCCACTCTGGACACATGGCTCCCAGCAG 240 |
| Dd | 347 | CTCGGGGTACACATATCGAAGTTTCGATCTGCCACTCTGGACACATGGCTCCCAGCAG 406 |
| Qy | 241 | GTTGCATTTATACAGTCAATGGGAAATGATAAAGCAAATAGTTACTCGGGAAGCAGACTA 300 |
| Dd | 407 | GTTGCATTTATACAGTCAATGGGAAATGATAAAGCAAATAGTTACTCGGGAAGCAGACTA 466 |
| Qy | 301 | CCCCAAAATATGATAGAGTTGGAAATTTGAGAAATTTTATACGTGCAAAAGTATGAGAGA 360 |
| Dd | 467 | CCCCAAAATATGATAGAGTTGGAAATTTGAGAAATTTTATACGTGCAAAAGTATGAGAGA 526 |
| Qy | 361 | AGATGGGTTTCTAGAGGGGAAAGCGCTAGATCACCTCTAGAGTCGAGCAGGAACCGCGG 420 |
| Dd | 527 | AGATGGGTTTCTAGAGGGGAAAGCGCTAGATCACCTCTAGAGTCGAGCAGGAAGCGCGG 586 |
| Qy | 421 | AAATCTGTGGAGAGAAAGTGGGCCCGGGATATGAGCATGGACATAGTAGTCTGTGTAAT 480 |
| Dd | 587 | AAATCTGTGGAGACAAAGTGGGCCCGGGATATGAGCATGGACATAGTAGTCTGTGTAAT 646 |
| Qy | 481 | TTGTTTGAGGAGAGAAACCTATTTCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 540 |
| Dd | 647 | TTGTTTGAGGAGAGAAACCTATTTCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 706 |
| Qy | 541 | AGAAATAATCTTCCCGTGCCTCCCAAGGACCAGTCAGGTATATAAGCCACACAGAGAA 600 |
| Dd | 707 | AGAAATAATCTTCCCGTGCCTCCCAAGGACCAGTCAGGTATATAAGCCACACAGAGAA 766 |
| Qy | 601 | ATGGAGTCTGCAGCTTACTCTCAGTAGAGGGGAGAAAAACAGCAGTAAATGTTGCAACGCA 660 |
| Dd | 767 | ATGGAGTCTGCAGCTTACTCTCAGTAGAGGGGAGAAAAACAGCAGTAAATGTTGCAACGCA 826 |
| Qy | 661 | TCAGATCTCTCAAAGGTGGATTTTGTCTACTGATCTGTTTTAACATGCTCTCAATGGATGAT 720 |
| Dd | 827 | TCAGATCTCTCAAAGGTGGATTTTGTCTACTGATCTGTTTTAACATGCTCTCAATGGATGAT 886 |
| Qy | 721 | TCGACTACAAATACCTCAGAGGCAACTCTCGGCCATCTCTCGCGGATGATAACTCATGG 780 |
| Dd | 887 | TCGACTACAAATACCTCAGAGGCAACTCTCGGCCATCTCTCGCGGATGATAACTCATGG 946 |
| Qy | 781 | GCTGGCTTTTCAGTCTGCTGGAAGTGGTCAAAACGGCAGAGAAAAATGTCACAGCCAAGCCT 840 |
| Dd | 947 | GCTGGCTTTTCAGTCTGCTGGAAGTGGTCAAAACGGCAGAGAAAAATGTCACAGCCAAGCCT 1006 |
| Qy | 841 | GCTGAGAGCAGTTCTCTCCAGCTTCACTTTCTGACTTTGAGGATTTGTTTAAAGGACACA 900 |
| Dd | 1007 | GCTGAGAGCAGTTCTCTCCCTCAGCTTCACTTTCTGACTTTGAGGATTTGTTTAAAGGACACA 1066 |
| Qy | 901 | CCTAATTTAAACAACTCAACAGGACCAAAAGATGTAAGGCGGATATCATGAGCCTGTTT 960 |

Db 1067 CCTAATTAACAACTCAACAGCACCAAAAGATGTGAAAGGCGATATCATGAGCCTGTTT 1126
Qy 961 GAGAAAGCAGNATAGTATCGCTTTTGGCATGCTATCAGCAGAGGTTGCTATGCTCGCT 1020
Db 1127 GAGAAAGCAGNATAGTATCGCTTTTGGCATGCTATCAGCAGAGGTTGCTATGCTCGCT 1186
Qy 1021 CAGCAGCAAGCCCTTTACATGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080
Db 1187 CAGCAGCAAGCCCTTTACATGCTGCGAGGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1246
Qy 1081 AATCAACAGAGTATGCTTAATGCTCTTAACGCTAGCTTTCTGCAAAATGGTCAAAACCCCGGC 1140
Db 1247 AATCAACAGAGTATGCTTAATGCTCTTAACGCTAGCTTTCTGCAAAATGGTCAAAACCCCGGC 1306
Qy 1141 GGTACACAGTCCCGGATGCTAATACCGGCTAGGCTGCTCAAGCTGATCTCCAGAACTT 1200
Db 1307 GGTACACAGTCCCGGATGCTAATACCGGCTAGGCTGCTCAAGCTGATCTCCAGAACTT 1366
Qy 1201 ATGCAAAACATGATATGAACGCAACATGAACACGAGACCGCAACCGCAAGAGAAC 1260
Db 1367 ATGCAAAACATGATATGAACGCAACATGAACACGAGACCGCAACCGCAAGAGAAC 1426
Qy 1261 ACTCTAATACCATCATCCAGTTCTTACAAATGGGCTCAAGCTAATCAAGTGAACGCT 1320
Db 1427 ACTCTAATACCATCATCCAGTTCTTACAAATGGGCTCAAGCTAATCAAGTGAACGCT 1486
Qy 1321 ATGACCCCAACTCAACCGGTAACTCAGTCAATCCGCAACCGCAACCGCAAGAGAAC 1380
Db 1487 ATGACCCCAACTCAACCGGTAACTCAGTCAATCCGCAACCGCAACCGCAAGAGAAC 1546
Qy 1381 ACACCATCTTCAAAATCAGCAAGAGCTTTGATTTCTTCTTGATGATGGAATGTTTC 1440
Db 1547 ACACCATCTTCAAAATCAGCAAGAGCTTTGATTTCTTCTTGATGATGGAATGTTTC 1606
Qy 1441 ACAAACATGTA 1452
Db 1607 ACAAACATGTA 1618

RESULT 7

ACN48712
ID ACN48712 standard; cDNA; 596 BP.
XX
AC ACN48712;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton primed seed EST Clone ID: LIB3825-026-Q6-K6-H9, SEQ:3493.
XX
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
KW variety DP508; library LIB3825; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS Gossypium hirsutum.
XX
PN US2004123340-A1.
XX
PD 24-JUN-2004.
XX
PF 12-DEC-2001; 2001US-00021323.
XX
PR 14-DEC-2000; 2000US-0255619P.
XX
XX (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
PI WPI; 2004-479808/45.
XX

XX
PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
PS Claim 1; SEQ ID NO 3493; 34pp; English.
XX
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC AKM45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP508, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoceium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucotton338. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety DP508 primed seed cDNA library (LIB3825). The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ Sequence 596 BP; 173 A; 127 C; 141 G; 154 T; 0 U; 1 Other;

Query Match 20.5%; Score 298.2; DB 13; Length 596;
Best Local Similarity 84.2%; Pred. No. 4.4e-80;
Matches 336; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
Qy 1 ATGAACGAGAAAGCAACGCTCTTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAA 60
Db 198 ATGAACGAGAAAGCAACGCTTACCAGGAGCTCAAGCCACACACCGAAGATTATTGNA 257
Qy 61 GGGCTTCTTAAACATCCAGAGAACAGAGAATGTCTGCTGCTCAAAACAAAAGGTCCAAGA 120
Db 258 AGTCTTCTTAAATTTGCCGAGAAATAGGAATGCGCGACTGCAAGCCAAAGGTCGAGA 317
Qy 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCGATGCAATGTTCTGGGATTCACAGGAT 180
Db 318 TGGGCAAGTGTGAATTTAGGTATCTTTATATGCAATGCAATGTTCTGGGATTCACAGGAT 377
Qy 181 CTCGGGTACACATATCGAAGGTTGATCTGCCACTCTGGACACATGCTCCCCGAGCAG 240
Db 378 CTTGGGGTACACATATCGAAGGTTGATCTGCTGCTACCTAGACACATGCTTCTGAGCAG 437
Qy 241 GTTGCAATTATACAGTCAATGGGAAATGATAAAGCAATAGTTTACTGGGAAGCAGAGCTA 300
Db 438 GTTGCTTTTATTTCAATCAATGGGAAATGATAAAGCAACAGTTACTGGGAAGCTGAGTTA 497
Qy 301 CCCCACAACTATGATAGAGTTGGAATTTGAGATTTTATACGTGCAAGTATGAGAGAG 360
Db 498 CCCCCTAATTTATGATAGAGTTGGAATTTGAGAACTTTCATCCGTGCAAGTATGAGAGAAA 557
Qy 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCT 399
Db 558 AGATGGGTTTCTAGAGATGGAATAATCCAAATCACCATCT 596

RESULT 8

AA48972/c
ID AAL48972 standard; DNA; 5580 BP.

XX AAL48972;
XX DT 28-OCT-2002 (first entry)
XX A thaliana NEVERSHED protein gene.
XX NEVERSHED; ARF; GAP; GTPase activating protein domain; chromosome 5;
KW GTP-binding protein domain; organ abscission; floral abscission; plant;
KW gene; db.

OS Arabidopsis thaliana.

XX WO200261042-A2.
XX PD 08-AUG-2002.

XX PF 22-JAN-2002; 2002WO-US001938.
XX PR 29-JAN-2001; 2001US-0264974P.

XX (SALK) SALK INST BIOLOGICAL STUDIES.
XX PA (REGC) UNIV CALIFORNIA.
XX Liljegren SJ, Ecker JR, Yanofsky MF;
PI WPI; 2002-627475/67.

XX New mutant plants comprising a modified ARF-GAP domain and having
PT decreased organ abscission, particularly floral abscission.
XX Example 2; Page 54-56; 67pp; English.

XX The present invention relates to a mutant plant which has decreased organ
CC abscission, comprising a nucleotide sequence having a modified ARF GAP
CC domain. The mutant plant is useful as a source of plants or progenies
CC having reduced or decreased floral abscission. The present sequence is
CC the Arabidopsis thaliana NEVERSHED protein coding sequence, which can be
CC mutated in plants of the invention

SX Sequence 5580 BP; 1821 A; 1010 C; 1034 G; 1715 T; 0 U; 0 Other;

Query Match 17.6%; Score 255.4; DB 6; Length 5580;
Best Local Similarity 60.1%; Pred. No. 1.8e-66;
Matches 877; Conservative 0; Mismatches 1; Indels 581; Gaps 4;
QY 575 GTCAGGTTATAAGCCACAGCAGAAATGGAGTCTGCAGTACTCCAGTAGAGGGGAGA 634
DB 1482 GACAGGTTATAAGCCACAGCAGAAATGGAGTCTGCAGTACTCCAGTAGAGGGGAGA 1423
QY 635 AACAGCAGTAAATGTTGACAGCAGATCAGATCCTCCAAAGTGGATTGTTCTACTGATC 694
DB 1422 AACAGCAGTAAATGTTGACAGCAGATCAGATCCTCCAAAGTGGATTGTTCTACTGATC 1363
QY 695 TGTAAACATGCTATCAATGGATGATTCGACTACAAATACCTCAGAGGCAACTCTGGCG 754
DB 1362 TGTAAACATGCTATCAATGGATGATTCGACTACAAATACCTCAGAGGCAACTCTGGCG 1303
QY 755 ATACTCCTCCGATGATACTCATGGCGTGGCTTCAGT----- 793
DB 1302 ATACTCCTCCGATGATACTCATGGCGTGGCTTCAGTGTATGATCTATCTAGCATGT 1243
QY 794 ----- 793
DB 1242 GATAATCTAACACATCGTGGCACATTTTTCCTATCCTTTGTTAACTTTATGTCATTTTCT 1183
QY 794 -----CTGCTGG 800
DB 1182 CTTTGAATAATGCTTTGGAGTAATAGTAATTTCTGAATCTTAAACTGTAGCTGCTGG 1123

QY 801 AAGTGGTCAAAACGCGAGAGAAAATTTGTCAAGCCAAAGCTGCTGAGAGCAGTTCTCTCC 860
DB 1122 AAGTGGTCAAAACGCGAGAGAAAATTTGTCAAGCCAAAGCTGCTGAGAGCAGTTCTCTCC 1063
QY 861 AGCTTCATCTTCTGACCTTTGAGGATTTGTTTAAAGGACACACCTAATTTAACTCAACA 920
DB 1062 AGCTTCATCTTCTGACCTTTGAGGATTTGTTTAAAGGACACACCTAATTTAACTCAACA 1003
QY 921 AGCACCAAAAGATGTGAAGGGGATATCATGAGCTGTTTGAGA----- 964
DB 1002 AGCACCAAAAGATGTGAAGGGGATATCATGAGCTGTTTGAGAAGGTAATTTATAGTT 943
QY 965 ----- 964
DB 942 AATTATATACGTTCTCAGATTTCACTGCTTTGTTGATATTCAAGCTTTTCCAAAGTTTCGTT 883
QY 965 ----- 964
DB 882 TTTGTATCATTATTAACTTACATTTTCAAAATATAGCGAATGATATGATTTGTGTTTGT 823
QY 965 ----- 964
DB 822 TATCATTACTATGATTTTCTCAACCGGACATTTGGAAACACGAGAGGATTTCAAGTG 763
QY 965 -----AGACGAATATAGTATCGCCTTTTGGCCATGCA 995
DB 762 ATATTAACCTCTGTGTCTTATTTGATTACAGAGATATAGTATCGCCTTTTGGCCATGCA 703
QY 996 TCAGCAACAGGTTGCTATGCTCGCTCAGCAGCAAGCCCTTTACATGGCTGCGAGCAAGC 1055
DB 702 TCAGCAACAGGTTGCTATGCTCGCTCAGCAGCAAGCCCTTTACATGGCTGCGAGCAAGC 643
QY 1056 TCCTGGAGGCACTCCAAACGGCGTGAATCAACAGCTATGCTTAATGCTCTTAACTGATG 1115
DB 642 TGCTGGAGGCACTCCAAACGGCGTGAATCAACAGCTATGCTTAATGCTCTTAACTGATG 583
QY 1116 TTCTGCAATTTGGTCAAAACCGCGGCTACAGATCCCGGAATGACTTAACCCCGTAGG 1175
DB 582 TTCTGCAATTTGGTCAAAACCGCGGCTACAGATCCCGGAATGACTTAACCCCGTAGG 523
QY 1176 TGCTCAAGCTGATCTCCAGAACTTATGCA----- 1206
DB 522 TGGTCAAGCTGATCTCCAGAACTTATGCAAGTATGCTTCCCTTTACTTTTAAAGC 463
QY 1207 ----- 1206
DB 462 AAGATCAATATTAACAATTGACCAAAAGTTCTTGCAAAATTTCTCACTAATTTCTGCTCTG 403
QY 1207 -----AACATGAATATGAACGCAAAACATGAACACGAGACCGGCAC 1246
DB 402 TTACATCCCATTAATTTTAGAACATGATATGAACGCAAAACATGAACACGAGACCGGCAC 343
QY 1247 AACCGCAAGAGAACACTCTACAAATACCATATC----- 1280
DB 342 AACCGCAAGAGAACACTCTACAAATACCATATC----- 283
QY 1281 ----- 1280
DB 282 GTTTGTTTCTTCTGAAACATCATAAAGTTATCAAAAGTCTTAACTTAACCTAATCTCA 223
QY 1281 -----CAGTTTCTACAAATGGGTCAAGCTTAATCAAGT 1313
DB 222 TATGATCTTTGGTGTGTTTATTAATGAGTTTCTACAAATGGGTCAAGCTTAATCAAGT 163
QY 1314 GAACGGTATGACCCCAAACTCAACCGTAAACCTCAGTCAATCATCGGCAACCCCAAC 1373
DB 162 GAACGGTATGACCCCAAACTCAACCGGTAAACCTCAGTCAATCATCGGCAACCCCAAC 103
QY 1374 AAGCAACCAACCATCTTTCACATCAGGCAAGACTTTGATTTCTTCTTCTGATGATGG 1433
DB 102 AAGCAACCAACCATCTTTCACATCAGGCAAGACTTTGATTTCTTCTTCTGATGATGG 43

```
QY 1434 AATGTTACAAACATTTGA 1452
DB |||||
42 AATGTTACAAACATTTGA 24

RESULT 9
AAC75657
ID AAC75657 standard; cDNA; 371 BP.
XX
XX AAC75657;
AC
XX
XX
DT 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF1212 polynucleotide sequence SEQ ID NO:2423.
DE
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
XX
XX 02-APR-1999; 99US-0127636P.
XX
XX 05-APR-1999; 99US-0127728P.
XX
XX 30-MAR-2000; 2000US-00540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX
XX P-PSDB; AAB41448.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 5; Page 1753-1754; 5507pp; English.
XX
XX AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancer, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
```

```
disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 371 BP; 113 A; 74 C; 102 G; 82 T; 0 U; 0 Other;
Query Match 16.3%; Score 237.2; DB 3; Length 371;
Best Local Similarity 81.4%; Pred. No. 1.5e-61;
Matches 275; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1 ATGAACGAGAAAGCCACGCTCTTAAGGAGCTTAATGCCGCCCATAGAAAGATTTCTTGA 60
DB |||||
34 ATGAACGAGAGGGCTCGTCTCAAGGAGCTCAACGCCAAGCACACAGAGATATTGGA 93
QY 61 GGGCTTCTTAACATCCAGAGAACAGAGAAATGCTGACTGCAAAACAAAGGTCCAAGA 120
DB |||||
94 GGTCTTCTACGGCATCTCTGAGAAATAGGAGACTGCGAGACTCAAGTCANAGGCTCTCGA 153
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATCTTCTGGGATTCACAGGAGT 180
DB |||||
154 TGGCAAGTGTGATCTAGGTATCTTTATATGATGATGATGATCTTCTGGCATTCATAGAAGC 213
QY 181 CTCGGGGTACACATATCAATGAGGTTTCGATCTGCCACTCTTGGACATATGCTCCCCAGAG 240
DB |||||
214 CTGGGGGTGCACATATCTAAGGTAAAGATCTGCCACCTCTGGATACATATGCTGCCAGAGCA 273
QY 241 GTTCATTTATACAGTCAATGGGAAATGATAAAGCAAAATAGTACTTGGGAAGCAGAGCTA 300
DB |||||
274 GTTGCAATTTATCAATCAATGGGAAACGAAAGCAAAATAGTACTTGGGAAGCAGAGCTG 333
QY 301 CCCCCAACTATGATAGAGTTGGAAATTGAGAAATTTAT 338
DB |||||
334 CCTCTAACTACGATAGGTTGGTAATAGAGAAATTTGAT 371

RESULT 10
ABN26360
ID ABN26360 standard; cDNA; 371 BP.
XX
XX AC ABN26360;
XX
XX 24-JUN-2002 (first entry)
XX
XX Human ORFX polynucleotide sequence SEQ ID NO:21197.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192523-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US010836.
XX
XX 30-MAY-2000; 2000US-0206132P.
XX
XX 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach MD;
XX
XX WPI; 2002-106308/14.
XX
XX P-PSDB; ABP10608.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
```


DB 241 GCCTCCTAACTAGTAGGTTGGAAATAGAGAAATTTTCATCCGTGCAAAATATGAGGACAA 300

QY 360 GAGATGGGT 368
|||||

DB 301 GAGATGGGT 309

RESULT 12
ABX82114
ID ABX82114 standard; cDNA; 270 BP.
XX
AC ABX82114;
XX
DT 24-APR-2003 (first entry)
XX
DE Corn ear-derived polynucleotide (cpd) #574.
XX
KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene transcription; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desired characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.
XX
OS Zea mays.
XX
PN US6476212-B1.
XX
PD 05-NOV-2002.
XX
PF 14-MAY-1999; 99US-00313294.
XX
PR 26-MAY-1998; 98US-0086722P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX WPI; 2003-2008840/20.
DR
XX Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with growth,
PT development.
XX
PS Example; SEQ ID NO 574; 390pp; English.
XX
CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC and SATMON023. Some of the cdps uniquely identify structural, functional,
CC and regulatory genes of corn ear. The polynucleotides sequences are
CC useful for detecting cdps in a sample, for producing a corn ear-specific
CC profile of gene transcription, for detecting altered gene expression in
CC inbred or hybrid plants, and for screening several molecules for specific
CC binding to the polynucleotide. The cdps are useful to identify, isolate,
CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridisation techniques
CC to follow the expression of desirable traits through plant breeding
CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth
CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cdps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cdps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to
CC detect or confirm conditions or diseases associated with abnormal levels
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC polynucleotides (cdps) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USFTO web site at
CC seqdata.uspto.gov/psipdsIDEntry.html
XX

SQ Sequence 270 BP; 85 A; 51 C; 70 G; 64 T; 0 U; 0 Other;

Query Match 13.7%; Score 199.6; DB 10; Length 270;
Best Local Similarity 83.7%; Pred. No. 4e-50;
Matches 226; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 92 GTGCTGACTGC AAAAAGGTC CAAAGATGGGCTAGTGT TAAATTTAGGTATCTTTATCT 151
DB 1 CGCGAGACTGCAAGTCAAAGGTCCTCGATGGGCAAGTGTGAATCTAGGTATCTTTATAT 60
QY 152 GCATGCATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCGAAGGTTGATCTG 211
DB 61 GCATGCATGTTCTGGCAATTCATAGAGCCCTGGGGGTGCAATATCTAAGGTAAGATCTG 120
QY 212 CCACCTCTGGACATATGGCTCCCGAGCAGGTTGCATTTATATACAGTCAATGGGAAATGATA 271
DB 121 CCACCTCTGGATACATGGCTGCCAGAGCAAGTTGATTTTCAATCAATGGGAAACGAA 180
QY 272 AAGCAATAGTTACTGGGAAAGCAGAGCTACCCCAAACTATGATAGAGTTGGAATTGAGA 331
DB 181 AAGCAATAGCTATTGGGAAAGCAGAGCTCCCTCTACTACGATAGGTTGGAAATAGAGA 240
QY 332 ATTTTATACGTGCAAGTATGAAGAGAGA 361
DB 241 ATTTTCATCCGTGCAAAATATATGAGGACAAGA 270

RESULT 13
AAC56661
ID AAC56661 standard; DNA; 408 BP.
XX
AC AAC56661;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #532.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS Eucalyptus grandis.
XX
PN WO2000053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US0006112.
XX
PR 11-MAR-1999; 99US-00266513.
XX 18-AUG-1999; 99US-0149485P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX WPI; 2000-579369/54.
DR
XX New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide.
XX
PS Claim 1; Page 486; 747pp; English.
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.

CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
 CC and CREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements and
 CC MYB
 XX
 SQ Sequence 408 BP; 105 A; 96 C; 109 G; 98 T; 0 U; 0 Other;
 Query Match 12.6%; Score 182.4; DB 3; Length 408;
 Best Local Similarity 85.0%; Pred. No. 9.1e-45;
 Matches 204; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 QY 1 ATGAACGAGAAAGCCACGCTCTTAAGGAGCTTAATCCCGGCATAGAAAGATCTTGAA 60
 DB 169 ATGAACGAGAAAGCCACGCTCTTAAGGAGCTTAATCCCGGCATAGAAAGATCTTGAA 228
 QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
 DB 229 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 288
 QY 121 TGGGCTAGTGTAAATTTAGGTATCTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
 DB 289 TGGGCTAGTGTCAATTTAGGATATTTATATGCAATGTTCTGGGATTCACAGAGT 348
 QY 181 CTGGGGTACACATATCGAAGTTCGATCTGCCACTCTGCACACATGGCTCCCGGAGCAG 240
 DB 349 CTGGGGTACACATATCGAAGTTCGATCTGCCACTCTGCACACATGGCTCCCGGAGCAG 408
 RESULT 14
 ABL73349
 ID ABL73349 standard; cdna; 290 BP.
 XX AC ABL73349;
 XX
 DT 14-MAY-2002 (first entry)
 XX
 DE Corn tassal-derived polynucleotide (cdps) SEQ ID NO:2723.
 XX
 KW Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPs;
 KW inheritance; characteristic; growth; development; disease resistance;
 KW environmental adaptability; quality; yield; molecular marker;
 KW multigene trait; plant breeding; corn tassal; gene; ss.
 XX
 OS Zea mays.
 XX
 XX US2001051335-A1.
 XX
 XX 13-DEC-2001.
 XX
 XX 16-APR-1999; 99US-00294093.
 XX
 XX 21-APR-1998; 98US-0082567P.
 XX
 XX (LALG/) LALGUDI R V.
 XX (ITOL/) ITO L Y.
 XX (SHER/) SHERMAN B K.
 XX
 XX Lalgudi RV, Ito LY, Sherman BK;
 XX
 XX WPI; 2002-163647/21.
 XX
 XX Novel purified corn tassal-derived polynucleotide useful for determining
 XX altered gene expression, to recover regulatory elements and to follow
 XX inheritance of desirable characteristics through hybrid breeding
 XX programs.
 XX
 XX Claim 1; SEQ ID NO 2723; 201pp; English.
 XX
 XX The present sequence describes a purified corn tassal-derived
 XX polynucleotide sequence (cdps) comprising a nucleic acid sequence
 XX selected from those given in ABL70627 to ABL76833. The cdps sequences

CC encode corn tassal-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful in
 CC the evaluation, and alteration of desired characteristics associated with
 CC growth and development, disease resistance, environmental adaptability,
 CC quality and yield, and as molecular markers for studying inheritance of
 CC multigene traits in a plant breeding program. (I) can be used to produce
 CC a tassal-specific profile of gene transcription, a transcript image, to
 CC clone regulatory elements for use in transformation vectors, to express a
 CC polypeptide, to identify, isolate or extend identical or related corn
 CC tassal nucleic acid sequences from DNA libraries, in nucleic acid
 CC hybridisation or amplification technologies, as query sequences to
 CC determine homology of known sequences, as probe for use in Southern or
 CC Northern hybridisation, and to identify the presence of and/or to
 CC determine the degree of similarity between two (or more) nucleic acid
 CC sequences
 XX
 SQ Sequence 290 BP; 83 A; 63 C; 66 G; 55 T; 0 U; 23 Other;
 Query Match 8.8%; Score 127.2; DB 6; Length 290;
 Best Local Similarity 75.9%; Pred. No. 5.7e-28;
 Matches 189; Conservative 0; Mismatches 57; Indels 3; Gaps 3;
 QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAGA 120
 DB 2 GGTCTTCTACGGCATCTCGAAGATAGAGAAATGCGCAGACTGCAAGTCAAGGGTCTCTCGA 61
 QY 121 TGGGCTAGTGTAAATTTAGGTATCTTATCTGCATGCAATGTTCTGG-GATTCACAGGAG 179
 DB 62 TGGGCAAGTGTGAATCTAGGTATCTTATATGATGATGATGTTCTGGACATTCATAGAAG 121
 QY 180 TCTCGGGGTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCA 239
 DB 122 CTGGGGGTGACATATCTAAGGTAGATCTGCCAC-CCGGATACATGCTGCCAGCA 180
 QY 240 GGTTCATTTATACAGTCAATCGGAATCATAAAGCAATAGTTACTGGGAAGCAGAGCT 299
 DB 181 AGTTGATTTTATTCATC-ANGGGACCCGAAANAGCAATCGNTATNGGNNACANACC 239
 QY 300 ACCCCCAAA 308
 DB 240 CCACCACACA 248
 RESULT 15
 ABQ98937
 ID ABQ98937 standard; DNA; 587 BP.
 XX AC ABQ98937;
 XX
 DT 04-NOV-2002 (first entry)
 XX
 DE Human ORF744 coding sequence.
 XX
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
 XX
 OS Homo sapiens.
 XX
 XX US2002082206-A1.
 XX
 XX 27-JUN-2002.
 XX
 XX 30-MAY-2001; 2001US-00867550.
 XX
 XX 30-MAY-2000; 2000US-0208427P.
 XX
 XX (LEAC/) LEACH M D.
 XX (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
FI WPI; 2002-626554/67.
XX P-PSDB; ABP64374.
DR
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
XX
PS Claim 2; SEQ ID NO 1487; 78pp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP64374-ABP64681 and ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ Sequence 587 BP; 166 A; 151 C; 138 G; 130 T; 0 U; 2 Other;

Query Match 7.4%; Score 107.2; DB 6; Length 587;
Best Local Similarity 60.1%; Pred. No. 1.1e-21;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 23 CTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAAGGGCTTCTTAAACATCCAGAGA 82
DB 119 CTCAGAAGCTGAACGAGCAGCAGCAGCTCATCTATCCAAGCTTCTGAGGGAGGAGCA 178
QY 83 ACAGAGAAATGCTGACTGCTGCAAAACAAAAGGTCCAGATGGGCTAGTGTATTTAGGTA 142
DB 179 ACAAGTACTGCGCCGACTCGAGGCCAAAGGTCTCTCGATGGGCTTCTCGAATATTTGGTG 238
QY 143 TCTTTATCTGCATGCAATGTTCTGGGATTACAGAGAGTCTCGGGGTACACATATCGAAGG 202
DB 239 TGTATTTTGCATCAGATGCTGGAAATTCATAGAAATCTTGGGTTTCATATATCCAGGG 298
QY 203 TTCGATCTGCCACTCTGGACACATGCTCCCGAGCAGGTTGCATTTATACAGTCAATGG 262
DB 299 TCAATCAGTCAACCTAGACCAATGACAGCAGACAGATACAGTGCATGCAAGATATGG 358
QY 263 GAATGATTAAGCAATAGTTACTTGGGAGCAGAGCTACCCCAACTATGATAGA 318
DB 359 GAAATATAAGCAAGCACTACTCTATGAAGCCAACTCTCCAGAGAACTTTGGAAGA 414

Search completed: September 27, 2005, 03:19:28
Job time : 785 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:40:34 ; Search time 6231 Seconds
(without alignments)
11291.442 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaagccaacgt.....gaatgttcacaaacattga 1452

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 1452 | 100.0 | 1452 | 6 | CQ803692 Sequence |
| 2 | 1452 | 100.0 | 1612 | 8 | BT000287 Arabidops |
| 3 | 1452 | 100.0 | 1777 | 8 | AY099716 Arabidops |
| 4 | 1447.2 | 99.7 | 1794 | 8 | AY088913 Arabidops |
| 5 | 1436 | 98.9 | 1492 | 8 | AF184144 Arabidops |
| 6 | 307 | 21.1 | 1451 | 8 | AK067240 Oryza sat |
| 7 | 307 | 21.1 | 1741 | 8 | AK067125 Oryza sat |
| 8 | 307 | 21.1 | 2157 | 8 | AK068892 Oryza sat |
| 9 | 255.4 | 17.6 | 78596 | 8 | AB010695 Arabidops |
| 10 | 237.2 | 16.3 | 371 | 6 | CQ455437 Sequence |
| 11 | 199.6 | 13.7 | 270 | 6 | AR245215 Sequence |
| 12 | 111.2 | 7.7 | 3504 | 5 | BC077937 Xenopus l |
| 13 | 109.6 | 7.5 | 778 | 5 | BC073437 Xenopus l |
| 14 | 107.2 | 7.4 | 554 | 9 | BC086672 Homo sapi |
| 15 | 107.2 | 7.4 | 2223 | 9 | AY055004 Homo sapi |
| 16 | 107.2 | 7.4 | 2304 | 9 | AY055003 Homo sapi |
| 17 | 107.2 | 7.4 | 2335 | 9 | BC028074 Homo sapi |
| 18 | 107.2 | 7.4 | 2885 | 6 | AR541755 Sequence |
| 19 | 107.2 | 7.4 | 3117 | 6 | AX059559 Sequence |

| | | | | | |
|----|-------|-----|--------|----|--------------------------------|
| 20 | 107.2 | 7.4 | 3164 | 9 | AF442495 |
| 21 | 107.2 | 7.4 | 3219 | 9 | BC036123 Homo sapi |
| 22 | 105.6 | 7.3 | 799 | 6 | BD148686 Homo sapi |
| 23 | 105.6 | 7.3 | 799 | 6 | AX868624 Sequence |
| 24 | 105.6 | 7.3 | 2314 | 6 | BD159879 Sequence |
| 25 | 105.6 | 7.3 | 2314 | 6 | AX882709 Sequence |
| 26 | 105.6 | 7.3 | 2314 | 9 | AK023221 Homo sapi |
| 27 | 101 | 7.0 | 2286 | 10 | BC006946 Mus muscu |
| 28 | 98.4 | 6.8 | 2916 | 10 | BC052413 Mus muscu |
| 29 | 98 | 6.7 | 3670 | 5 | BC074142 Xenopus l |
| 30 | 95.6 | 6.6 | 1540 | 8 | AK059690 Oryza sat |
| 31 | 94.8 | 6.5 | 2561 | 6 | AX059554 Sequence |
| 32 | 94.4 | 6.5 | 705 | 5 | BC071454 Danio rer |
| 33 | 94.2 | 6.5 | 1447 | 6 | AR220880 Sequence |
| 34 | 94.2 | 6.5 | 1976 | 6 | AR339024 Sequence |
| 35 | 94.2 | 6.5 | 2457 | 9 | HS228H131 Novel hum |
| 36 | 94.2 | 6.5 | 2618 | 9 | BC021133 Homo sapi |
| 37 | 94 | 6.5 | 110000 | 8 | CR382134_07 Continuation (8 of |
| 38 | 92.2 | 6.3 | 2541 | 6 | BD194871 86 human |
| 39 | 92.2 | 6.3 | 2541 | 6 | CQ855199 Sequence |
| 40 | 89.4 | 6.2 | 340552 | 3 | PFA929154 Plasmodi |
| 41 | 88.6 | 6.1 | 1933 | 5 | BC060484 Xenopus l |
| 42 | 86.2 | 5.9 | 309 | 6 | AX886856 Sequence |
| 43 | 86.2 | 5.9 | 309 | 6 | BD026466 Sequence |
| 44 | 86.2 | 5.9 | 459 | 6 | AR413208 Sequence |
| 45 | 86.2 | 5.9 | 459 | 6 | AX970042 Sequence |

ALIGNMENTS

RESULT 1
CQ803692
LOCUS CQ803692 1452 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 103 from Patent WO2004035798.
ACCESSION CQ803692
VERSION CQ803692.1 GI:47110480
KEYWORDS Arabidopsis thaliana (chale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 103 29-APR-2004;
CropDesign N.V. (BE)

FEATURES
source
1..1452
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 100.0%; Score 1452; DB 6; Length 1452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACGAGAAAGCCAAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 60
DB 1 ATGAACGAGAAAGCCAAACGCTCTTAAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 60
QY 61 GGGCTTCTTAACATCCAGAGAACAGAGATGTGTGACTGCANAAACAAAGGTCCAAGA 120
DB 61 GGGCTTCTTAACATCCAGAGAACAGAGATGTGTGACTGCANAAACAAAGGTCCAAGA 120
QY 121 TCGGCTAGTGGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
DB 121 TCGGCTAGTGGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
QY 181 CTCGGGGTACACATATCGAAGGTTCCATCTGCACCTCTCGACACATCGGTCTCCCGAGCAG 240

```

Db      181  CTGGGGTACACATATCGAAGGTTGCGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
Qy      241  GTTGCATTTATACAGTCAATGGGAATGATAAGCAATATAGTTACTTGGGAACGAGCTA 300
Db      241  GTTGCATTTATACAGTCAATGGGAATGATAAGCAATATAGTTACTTGGGAACGAGCTA 300
Qy      301  CCCCACAACTATGATAGAGTGGAAATGGAATTTTATACGTGCAAGATGATGAAGAGAG 360
Db      301  CCCCACAACTATGATAGAGTGGAAATGGAATTTTATACGTGCAAGATGATGAAGAGAG 360
Qy      361  AGATGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTTAGAGTCGACGAGAAACGGCGG 420
Db      361  AGATGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTTAGAGTCGACGAGAAACGGCGG 420
Qy      421  AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGCATAGTAGTAGTCTCTTAAT 480
Db      421  AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATAGCATAGTAGTAGTCTCTTAAT 480
Qy      481  TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGAAACAAAGAAATAATGTTGCTCAACG 540
Db      481  TTGTTTGAGGAGAGAAACTATTCCAGCATCTAGAAACAAAGAAATAATGTTGCTCAACG 540
Qy      541  AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTCCAGTTTAAAGCCACAGCAGAAA 600
Db      541  AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTCCAGTTTAAAGCCACAGCAGAAA 600
Qy      601  ATGGAGTCTGAGCTACTCCAGTAGAGGGGAGAACAGCAGTAGTAATGTTGCACGCA 660
Db      601  ATGGAGTCTGAGCTACTCCAGTAGAGGGGAGAACAGCAGTAGTAATGTTGCACGCA 660
Qy      661  TCAGATCTCTCAAAAGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
Db      661  TCAGATCTCTCAAAAGTGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
Qy      721  TCGACTACAAATPACTCAGAGCAACTCTGCGCATCTCTGCGCATGATAACTCATGG 780
Db      721  TCGACTACAAATPACTCAGAGCAACTCTGCGCATCTCTGCGCATGATAACTCATGG 780
Qy      781  GTGGCTTTCAGTCTGCTGGAAGTGGTCAACCGCAGAGAAATGTCACGCCAAGCCT 840
Db      781  GTGGCTTTCAGTCTGCTGGAAGTGGTCAACCGCAGAGAAATGTCACGCCAAGCCT 840
Qy      841  GCTGAGAGCAGTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTAAGACACA 900
Db      841  GCTGAGAGCAGTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTAAGACACA 900
Qy      901  CCTAATTTAACTCAACAGCAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
Db      901  CCTAATTTAACTCAACAGCAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
Qy      961  GAGAGAGCAATATAGTATCGCTTTTGGCATGCTATCAGACAGGTTGCTATGCTCGCT 1020
Db      961  GAGAGAGCAATATAGTATCGCTTTTGGCATGCTATCAGACAGGTTGCTATGCTCGCT 1020
Qy      1021  CAGCAGCAAGCCCTTTACATGCTGTCAGCGAAAGCTGTGGAGGCACTCCAAACGGCGTG 1080
Db      1021  CAGCAGCAAGCCCTTTACATGCTGTCAGCGAAAGCTGTGGAGGCACTCCAAACGGCGTG 1080
Qy      1081  AATCAACAGCTATTGCTTAATGCTCTTAACTGATCTTCTGCAAAATGGTCAAAACCCCGC 1140
Db      1081  AATCAACAGCTATTGCTTAATGCTCTTAACTGATCTTCTGCAAAATGGTCAAAACCCCGC 1140
Qy      1141  GGCTACCAAGTCCCGGAATGATCAACCCGTAGGTGGTCAAGCTGATCTCCAGAACTT 1200
Db      1141  GGCTACCAAGTCCCGGAATGATCAACCCGTAGGTGGTCAAGCTGATCTCCAGAACTT 1200
Qy      1201  ATGCAAAACATCAATATGAACCAACATGAACACGAGACCCGCAACCCGCAAGAGAAC 1260
Db      1201  ATGCAAAACATCAATATGAACCAACATGAACACGAGACCCGCAACCCGCAAGAGAAC 1260
Qy      1261  ACTCTACAATACCCATCATCCAGTTTCTTACACAATGGGTCAAGCTTAATCAAGTGAACGGT 1320
Db      1261  ACTCTACAATACCCATCATCCAGTTTCTTACACAATGGGTCAAGCTTAATCAAGTGAACGGT 1320

```

```

Qy      1321  ATGACCCCAACTCAACGGGTAAACCTCAGTCATCATCCGCAACCCCAACCAACGACACC 1380
Db      1321  ATGACCCCAACTCAACGGGTAAACCTCAGTCATCATCCGCAACCCCAACCAACGACACC 1380
Qy      1381  ACACCATCTTCCAAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTC 1440
Db      1381  ACACCATCTTCCAAATCAGGCAAGACTTTGATTTCTCTCTTGTGATGATGGAATGTC 1440
Qy      1441  ACAAACAATTGA 1452
Db      1441  ACAAACAATTGA 1452

RESULT 2
BT000287      1612 bp      mRNA      linear      PLN 19-SEP-2002
LOCUS      Arabidopsis thaliana unknown protein (At5g54310) mRNA, complete cds.
DEFINITION
ACCESSION      BT000287      GI:23198157
VERSION      BT000287.1
KEYWORDS      FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE      1 (bases 1 to 1612)
AUTHORS      Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
                Palm,C.J., Chung,M.K., Hayashizaki,Y., Ishida,J., Chen,H.,
                Cheuk,R., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
                Kawai,J., Sakurai,T., Sacou,M., Seki,M., Shinn,P., Yamada,K.,
                Shinzaki,K., Becker,J., Theologis,A. and Davis,R.W.
                Direct Submission
TITLE      Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
JOURNAL      Stanford University, 855 California Avenue, Palo Alto, CA 94304,
                USA
COMMENT      e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Sacou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinzaki,K.

The Salk, Stanford, PCEC (SSP) Consortium members constructed and
sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M.,
Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,
Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M.,
Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,
Becker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinzaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
            Location/Qualifiers
            1..1612
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /db_xref="taxon:3702"
               /chromosome="5"
               /clone="U25083"
               /ecotype="Columbia"
               /notes="This clone is in pENTR/SD-dTopo This is a cloned
               PCR product using RIKEN clone RAFL09-80-M13 (AY099716) as
               a template"
            1..1612
               /gene="At5g54310"
            1..1452
               /gene="At5g54310"
               /codon_start=1

```

```
/product="unknown protein"
/protein_id="AANI5406.1"
/db_xref="GI:23198158"
/translation="MNEKANVSKELNARHKILEGLLKHPEENRECHADCKTKGPRNASV
NLGIFCMCGSHRSLGHSIKVRSATLDTLWPEQVFIQSMGNDKANSYWEALPP
NYDRVGIEFIKRAKYEKRWVSGEKARPPRVQEKRSVERSGYEHGSHSPVN
LPERKTIIPASRRNNVAATRLNLPVPOGPSOVIKPOOKMESAAATPVEREKOAVNVA
PASDPKVPDADLFLNLSMDSTTNTSATPGDTPADDNSWAGFOSGSGTAEKIV
TAKPSSSPASSPDEFLFKDTNLTQQAPKDKVGDIMSLFEKTNLIVSPFMHQQ
QVAMLAQOQALYNAAKAAAGGTGPNVQALNALNANWSNPNQGYQIPGHTNPVG
GQADLQKQNMNNMNMNTRPAQPOENTLOPSSSFYTMGQANOVNGMTNSTGKPK
SSSATOPTSTTPSSQSGKDFDFSLMDGMFKH"

ORIGIN
Query Match      100.0%; Score 1452; DB 8; Length 1612;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAACCAACGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
DB 1 ATGACGAGAAACCAACGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120
DB 61 GGGCTTTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAAACAAAGGTCCAAGA 120
QY 121 TGGGCTAGTGTAAATTAGGTATCTTTATCTGATGCAATGTTCTGGGATTTCACAGAGT 180
DB 121 TGGGCTAGTGTAAATTAGGTATCTTTATCTGATGCAATGTTCTGGGATTTCACAGAGT 180
QY 181 CTCGGGGTACACATATCGAAGTTCCATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240
DB 181 CTCGGGGTACACATATCGAAGTTCCATCTGCCACTCTGGACACATGGCTCCCGGAGCAG 240
QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
QY 301 CCCCCAAACTATGATAGAGTTGAATGAGAAATTTATACGTGCCAAAGATATGAGAGAG 360
DB 301 CCCCCAAACTATGATAGAGTTGAATGAGAAATTTATACGTGCCAAAGATATGAGAGAG 360
QY 361 AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTCAGAGCAAGAAACGGCG 420
DB 361 AGATGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTCAGAGCAAGAAACGGCG 420
QY 421 AATCTGTGAGAGAAAGTGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480
DB 421 AATCTGTGAGAGAAAGTGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480
QY 481 TTGTTTGAGAGAGGAAAACCTATTCAGCATCTAGAACTAGAACTAATGTTGCTGCAACG 540
DB 481 TTGTTTGAGAGAGGAAAACCTATTCAGCATCTAGAACTAGAACTAATGTTGCTGCAACG 540
QY 541 AGAATAAATCTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
DB 541 AGAATAAATCTCCCGTCTCCCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAA 600
QY 601 ATGGAGTCTCAGCTACTCCAGTAGAGAGGAGAAACAGCACTAATGTTGCAACGCA 660
DB 601 ATGGAGTCTCAGCTACTCCAGTAGAGAGGAGAAACAGCACTAATGTTGCAACGCA 660
QY 661 TCAGATCTCTCAAAGTGGATTTTGTAATGATCTGTTTAACTGTTAACTCAATGATGAT 720
DB 661 TCAGATCTCTCAAAGTGGATTTTGTAATGATCTGTTTAACTGTTAACTCAATGATGAT 720
QY 721 TCGACTCAAAATACCTCAGAGGCAACTCTCCGCGGATCTCTCCGATGATACTCATGG 780
DB 721 TCGACTCAAAATACCTCAGAGGCAACTCTCCGCGGATCTCTCCGATGATACTCATGG 780
QY 781 GCTGGCTTTCAGTCTGCTGGAAGTGGTCAACCGGACAGAGAAAATTTGTACAGCCAGACCT 840
DB 781 GCTGGCTTTCAGTCTGCTGGAAGTGGTCAACCGGACAGAGAAAATTTGTACAGCCAGACCT 840

RESULT 3
AY099716 1777 bp mRNA linear PLN 06-MAY-2002
LOCUS Arabidopsis thaliana unknown protein (At5g54310) mRNA, complete
DEFINITION cds
ACCESSION AY099716
VERSION AY099716.1 GI:20466499
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1777)
Nguyen.M., Karlin-Neumann.G., Southwick.A., Lam.B., Miranda.M.,
Palm.C.J., Bowser.L., Jones.T., Banh.J., Carninci.P., Chen.H.,
Cheuk.R., Chung.M.K., Hayaishizaki.Y., Ishida.J., Kamiya.A.,
Kawai.J., Kim.C., Lin.J., Liu.S.X., Narusaka.M., Pham.P.K.,
Sakano.H., Sakurai.T., Satou.M., Seki.M., Shinn.P., Yamada.K.,
Shinozaki.K., Ecker.J., Theologis.A. and Davis.R.W.
Direct Submission
Submitted (24-APR-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu
RIKEN Genomic Sciences Center (GSC) members carried out the
```

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Nguyen, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. 1777

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAFL09-80-M13 (R25083)"

/ecotype="Columbia"

/notes="This clone is in pBluescript"

1. 1777

/gene="At5g54310"

151. 1602

/gene="At5g54310"

/codon_start=1

/product="unknown protein"

/protein_id="AAM20567.1"

/db_xref="GI:20466500"

/translations="MNEKANVSKELNARHKILGKHPENRECAADCKTKGRNASV
NLGIFLCMCSGHRISLVHSKVRATLDTWLPQOVAFIQSNKNDKANSYWEALPP
NIDRVGIENIRAKIEKRVSRGEARSPRVEQRKRSVRSRGVYEHGSSSPVN
LFEERTIIPASRTNNVAATRLNLPVPPQSQVIRPQOMESAAATPVEREKQAVNVA
PASDPKVPFATDLFNLMDSDTTNTSEATPDTPADNSWAGSFQSGGTAEKIV
TAKPASSSPPASDFDLFKDTPNLTQQAPKQVKGDMISLFEKTNIVSPFAKHQ
QVAMLAQQALVMAAAAGGTPNGVQQAIALANLVASANNPNPGYQIPGWTNPVG
GQADLKLMNNMNMNTRPAQPOENTLOYPSSSFYTMGQANQVNGMTPNSTGPKPQ
SSSATQPTSTTPSSQSGKDFPFSSLMDFMTH"

ORIGIN

Query Match 100.0%; Score 1452; DB 8; Length 1777;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAAGCAACGCTCTTAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 60
DB 151 ATGAACGAGAAAGCAACGCTCTTAGGAGCTTAATGCCCGCCATAGAAGATTCTTGAA 210

QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGGTCCAAGA 120
DB 211 GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGGTCCAAGA 270

QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGACT 180
DB 271 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGACT 330

QY 181 CTGGGGTACACATATCGAAGGTTCGATCTGCACCTCTGACACATCGCTCCCCGAGCAG 240
DB 331 CTGGGGTACACATATCGAAGGTTCGATCTGCACCTCTGACACATCGCTCCCCGAGCAG 390

QY 241 GTTGCATTTATACAGTCATGGAATGTAAGCAATAGTTACTGGGAAGCAGAGCTA 300
DB 391 GTTGCATTTATACAGTCATGGAATGTAAGCAATAGTTACTGGGAAGCAGAGCTA 450

QY 301 CCCCCAACTATGATAGGTGGAAATTTAGAAATTTATACGTGCAAAAGTATGAAGAGAAG 360
DB 451 CCCCCAACTATGATAGGTGGAAATTTAGAAATTTATACGTGCAAAAGTATGAAGAGAAG 510

QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTTAGCTCAGCAGGAACGGCGG 420
DB 1591 AAAAAACATTGA 1602

DB 511 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTTAGAGTCGAGCAGGAACGGCGG 570
QY 421 AATCTCTGTGGAGAGAGTGGCGCGGATATAGCATGAGCATAGTGTAGTCTCTGTAAT 480
DB 571 AATCTCTGTGGAGAGAGTGGCGCGGATATAGCATGAGCATAGTGTAGTCTCTGTAAT 630
QY 481 TTGTTTGGAGGAGGAAACTATTCCAGCATCTAGAACAAAGAAATAAATTTGCTGCAACG 540
DB 631 TTGTTTGGAGGAGGAAACTATTCCAGCATCTAGAACAAAGAAATAAATTTGCTGCAACG 690
QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCTAGGTTTATAAGCCACAGCAGAA 600
DB 691 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCTAGGTTTATAAGCCACAGCAGAA 750
QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAAGCAGTAAATTTGTCACACGCA 660
DB 751 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAAGCAGTAAATTTGTCACACGCA 810
QY 661 TCAGATCTCCAAAGGTGGATTTTCTGTAATGCTGTTTAAACATGCTATCAATGGATGAT 720
DB 811 TCAGATCTCCAAAGGTGGATTTTCTGTAATGCTGTTTAAACATGCTATCAATGGATGAT 870
QY 721 TCGACTACAAATACCTCAGAGGCAACTCTCGCGGATCTCTCGCGGATGATAACTCATCG 780
DB 871 TCGACTACAAATACCTCAGAGGCAACTCTCGCGGATCTCTCGCGGATGATAACTCATCG 930
QY 781 GCTGGCTTTCACTCTGCTGGAGTGGTCAAAACGCGCAGAGAAATTTGTCAACCCAGCCT 840
DB 931 GCTGGCTTTCACTCTGCTGGAGTGGTCAAAACGCGCAGAGAAATTTGTCAACCCAGCCT 990
QY 841 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTTGGAGGATTTGTTTAAAGACACA 900
DB 991 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTTGGAGGATTTGTTTAAAGACACA 1050
QY 901 CCTAATTTAAACAATCAACAGCACAACAGATGTGNAAGGCGATATCATGAGCCCTGTTT 960
DB 1051 CCTAATTTAAACAATCAACAGCACAACAGATGTGNAAGGCGATATCATGAGCCCTGTTT 1110
QY 961 GAGAGAGCAATATAGTATCGCTTTTCCCATGCAATCAGCAACAGTTGCTATGCTCGCT 1020
DB 1111 GAGAGAGCAATATAGTATCGCTTTTCCCATGCAATCAGCAACAGTTGCTATGCTCGCT 1170
QY 1021 CAGCAGAAAGCCCTTTAATGCTGTCAGCGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080
DB 1171 CAGCAGAAAGCCCTTTAATGCTGTCAGCGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1230
QY 1081 AATCAACAGCTATTGCTTAATGCTTAAACGTAGCTTCTGCAAAATTTGGTCAAAACCGCGG 1140
DB 1231 AATCAACAGCTATTGCTTAATGCTTAAACGTAGCTTCTGCAAAATTTGGTCAAAACCGCGG 1290
QY 1141 GGCTACAGATCCCGGAAATGACTAAACCCGCTAGTGTGCTCAAGTGTCTCAGAAACTTT 1200
DB 1291 GGCTACAGATCCCGGAAATGACTAAACCCGCTAGTGTGCTCAAGTGTCTCAGAAACTTT 1350
QY 1201 ATGCAAAACATGAATATGAACGCAACATGAACAGAGACCCGCAACCGCAAGAGAC 1260
DB 1351 ATGCAAAACATGAATATGAACGCAACATGAACAGAGACCCGCAACCGCAAGAGAC 1410
QY 1261 ACTCTACAAATCCCATCATCATGTTTCTACAAATGGGTCAAGCTTAATCAAGTGAACCGT 1320
DB 1411 ACTCTACAAATCCCATCATCATGTTTCTACAAATGGGTCAAGCTTAATCAAGTGAACCGT 1470
QY 1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCTCATCATCCGCAACCCAAACAAAGCACC 1380
DB 1471 ATGACCCCAAACTCAACCGGTAAACCTCAGTCTCATCATCCGCAACCCAAACAAAGCACC 1530
QY 1381 ACACCATCTTCAATCAGGCAAGACTTGTGATTTCTCTTCTTCTGATGATGATGATGAT 1440
DB 1531 ACACCATCTTCAATCAGGCAAGACTTGTGATTTCTCTTCTTCTTCTGATGATGATGATGAT 1590
QY 1441 ACACCATCTTCA 1452
DB 1591 AAAAAACATTGA 1602

```

RESULT 4
AY088913 LOCUS 1794 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 99519 mRNA, complete sequence.
ACCESSION AY088913
VERSION AY088913.1 GI:21407687
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1794)
Hae, B.J., Volfovsky, N., Town, C.D., Troughan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
22088475
12093376
2 (bases 1 to 1794)
Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1794)
Brover, V., Troughan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source            1..1794
                        /organism="Arabidopsis thaliana"
                        /mol_type="mRNA"
                        /db_xref="taxon:3702"
                        /clone="99519"
                        167..1618
     CDS               /codon_start=1
                        /product="ARF GAP-like zinc finger-containing protein
                        ZIG3"
                        /protein_id="AA667219.1"
                        /db_xref="GI:21618169"
                        /translation="MNEKANVSKELNARHKILLEGLLKHPENRECAADCKTKGPRWASV
                        NIGIFCMCSGHRSLGVHISKVRSATLDTWLPEOVAFIQSMGNDKANSYWEAELEPP
                        NYDRVGIENFIRAKYIEKKRWKSGEKARSPRVQERRKSVETSGPYEHHSSSPVN
                        LFEERTIIPATFRNNVAATRLNLPVPPQSPQVVIKPPQKMSAATPVEREQVANVA
                        PASDPKVPFATDLFNLGMDSDTTNTSEATPDGTPADDSNAGSGQTAEKIV
                        TAKPASSSPPASSDFELFKDTPLNTTQQAPKDKVGDIMSIFEXTNIVSPFAMHQ
                        OVAMLAQQALYMAAKAFKGTGPNVQQAIALNALVANASNPNPGYVQIPGMTNPGQ
                        GQADLQKLMNNMNMNTRPAQPDENTLOVPSSSFYTMGQANQVNGWTPNSTGKPQ
                        SSSATQPTSTTSSQSGKQDFDSSLMDGMFTKH"
Query Match      99.7%; Score 1447.2; DB 8; Length 1794;
ORIGIN

```

```

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1449; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGACGAGAAAGCCAAACGCTCTTAAGGAGCTTAAATGCCGCATAGAAAGATTCTTTGAA 60
DB 167 ATGACGAGAAAGCCAAATGTCTTAAGGAGCTTAAATGCCGCATAGAAAGATTCTTTGAA 226
QY 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGATCTGCAAAACAAAAGGTCCAAGA 120
DB 227 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGTGTGATCTGCAAAACAAAAGGTCCAAGA 286
QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGACT 180
DB 287 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGACT 346
QY 181 CTGCGGGTACACATATCCAAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGACGAG 240
DB 347 CTCGGGGTACACATATCCAAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGGACGAG 406
QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
DB 407 GTTGCAATTTATACAGTCAATGGGAAATGATAAAGCAAAATAGTTACTGGGAAGCAGAGCTA 466
QY 301 CCCCACAACTATAGTAGTTGGAAATTCAGAAATTTTATACGTCGAAAGTATCAAGAGAGAG 360
DB 467 CCCCACAACTATAGTAGTTGGAAATTCAGAAATTTTATACGTCGAAAGTATCAAGAGAGAG 526
QY 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTTAGAGTCCAGCAGGAAACGGCGG 420
DB 527 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACCTCTTAGAGTCCAGCAGGAAACGGCGG 586
QY 421 AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGGAATAGTAGTAGTCTCTGTAAT 480
DB 587 AAATCTGTGGAGAGAAAGTGGGCGGGATATGAGCATGGAATAGTAGTAGTCTCTGTAAT 646
QY 481 TTGTTTGGAGAGAGAAACCTATTTCACGATCTAGAACAGAAATATGTTCTCTCAACG 540
DB 647 TTGTTTGGAGAGAGAAACCTATTTCACGATCTAGAACAGAAATATGTTCTCTCAACG 706
QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGACCCAGTCAAGTTTATAAGCCACAGCAGAGAA 600
DB 707 AGAATAAATCTTCCCGTCCCTCCCAAGACCCAGTCAAGTTTATAAGCCACAGCAGAGAA 766
QY 601 ATGAGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAGCAGTAATTTGCACGACA 660
DB 767 ATGAGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAGCAGTAATTTGCACGACA 826
QY 661 TCAGATCTCTCAAGGTGGATTTTCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
DB 827 TCAGATCTCTCAAGGTGGATTTTCTACTGATCTGTTTAAACATGCTATCAATGATGAT 886
QY 721 TCGACTTACAAATACCTCAGAGGCAACTCTCTGGCGATATCTCTGCCGATGATAAATCATGG 780
DB 887 TCGACTTACAAATACCTCAGAGGCAACTCTCTGGCGATATCTCTGCCGATGATAAATCATGG 946
QY 781 GCTGGCTTTTCAAGTCTGTGGAAGTGGTCAACCGGAGAGAAAATTTGTCAAGCAAGCCT 840
DB 947 GCTGGCTTTTCAAGTCTGTGGAAGTGGTCAACCGGAGAGAAAATTTGTCAAGCAAGCCT 1006
QY 841 GCTGAGAGAGTTCCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGGACACA 900
DB 1007 GCTGAGAGAGTTCCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGGACACA 1066
QY 901 CCTAATTTAAACAACTCAACAGCACAAAGATGTGAAAGGCGATATCATAGCCCTGTTT 960
DB 1067 CCTAATTTAAACAACTCAACAGCACAAAGATGTGAAAGGCGATATCATAGCCCTGTTT 1126
QY 961 GAGAAAGCAATATAGTATCGCCTTTTGGCATGCAATGCAACAGGTTGCTATGCTCGCT 1020
DB 1127 GAGAAAGCAATATAGTATCGCCTTTTGGCATGCAATGCAACAGGTTGCTATGCTCGCT 1186
QY 1021 CAGCAGAGAGCCCTTTATAGTGTGCTGAGGAAAGCTCTGGAGGCACTCCAAACGGCGTG 1080

```


Db 1187 CAGCAGCAAGCCCTTTACATGGCTGCAGCGAAGAGCTGCTGGAGCACTCCAAAACGGCGTG 1246

Qy 1081 ATCAACACAGCTATTGCTAATGCTCTTAAGCTAGCTTCTGCAAAATGGTCAACCCCGC 1140

Db 1247 AATCAACACAGCTATTGCTAATGCTCTTAAGCTAGCTTCTGCAAAATGGTCAACCCCGC 1306

Qy 1141 GGTACACAGATCCCGGAATGACTAACCCTAGTGGTCAAGCTGATCTCCAGAAACTT 1200

Db 1307 GGTACACAGATCCCGGAATGACTAACCCTAGTGGTCAAGCTGATCTCCAGAAACTT 1366

Qy 1201 ATGCAAAAATGATATGATGACGCAACATGACACGAGACCCGCAACCGCAAGAAC 1260

Db 1367 ATGCAAAAATGATATGATGACGCAACATGACACGAGACCCGCAACCGCAAGAAC 1426

Qy 1261 ACTCTACAATACCCATCATCTCAAGTCTTACAAATGGTCAAGCTAATCAAGTGAAGCT 1320

Db 1427 ACTCTACAATACCCATCATCTCAAGTCTTACAAATGGTCAAGCTAATCAAGTGAAGCT 1486

Qy 1321 ATGACCCCAAACTCAACCGGTAACTCAGTCAATCATCCGCAACCCCAACCAAGCAC 1380

Db 1487 ATGACCCCAAACTCAACCGGTAACTCAGTCAATCATCCGCAACCCCAACCAAGCAC 1546

Qy 1381 ACACATCTTCAATCAGCGCAAGACTTTGATTTCTCTTCTGATGATGGAATGTTTC 1440

Db 1547 ACACATCTTCAATCAGCGCAAGACTTTGATTTCTCTTCTGATGATGGAATGTTTC 1606

Qy 1441 ACAAAACATTGA 1452

Db 1607 ACAAAACATTGA 1618

RESULT 5

AF184144
LOCUS Arabidopsis thaliana 1492 bp mRNA linear PLN 19-NOV-2001
DEFINITION ZIG3 mRNA, partial cds.

ACCESSION AF184144
VERSION AF184144.1 GI:10441351

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1492)
Jensen,R.B., Lykke-Andersen,K., Frandsen,G.I., Nielsen,H.B., Haseloff,J., Jespersen,H.M., Mundy,J. and Skriver,K.
Promiscuous and specific phospholipid binding by domains in ZAC, a membrane-associated Arabidopsis protein with an ARF GAP zinc finger and a C2 domain
Plant Mol. Biol. 44 (6), 799-814 (2000)

JOURNAL 21070549
MEDLINE 11202441
PUBMED 2 (bases 1 to 1492)
REFERENCE Andersen,K.L., Jensen,R.B., Jespersen,H.M. and Skriver,K.
AUTHORS Direct Submission
TITLE Submitted (10-SEP-1999) Dept. of Protein Chemistry, Inst. of Molecular Biology, University of Copenhagen, Oe. Farimagsgade 2A, Copenhagen 1353, Denmark

FEATURES
source
1..1492
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
..1492
/note="similar to protein encoded by GenBank Accession Number AB010693"
/codon_start=3
/product="ARF GAP-like zinc finger-containing protein ZIG3"
/protein_id="AA017004.1"

CDS

/db_xref="GI:10441352"
/translation="VSKELNARHRKILLEGILKHPENRECADCKTKGPRWASVNLGIFI
CMQSGIHRISGVHISKVRSATLDTLWLPQVAFIQSMGNDRANSYWEALPENYDRVG
IENIRAKYEKRWVSERGEKASPRVEQERKSVESRGVGEHSSSPNLPFERK
TIPASRTNNVAATRLNLPVPPQATPQVITKQKMSAATPVERSKQAVNVAADPP
KVDPATDLFNMLSDMTTNTSEATPGDTPADDNSWAGFSAGSGOTAEKIVTAKPAE
SSQALYMAAKAAGCTGNGVNOQAIANALVANASWNPNGYQIIPGMTPVGGQADLQ
KLQNMNNANMNTTRPAQPOENTLOYPSSPYTMGOANOVNGMTPNSTKGKSSSATO
PTSTTPSSQSGKDFDFSLMDMGMFTKH"

ORIGIN

Query Match 98.9%; Score 1436; DB 8; Length 1492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ACCTCTCTAAGAGCTTAATGCCGCCCATAGAAAGATCTTGAAGGGCTCTTAAACATC 76
Db 1 ACCTCTCTAAGAGCTTAATGCCGCCCATAGAAAGATCTTGAAGGGCTCTTAAACATC 60

Qy 77 CAGAGAACAGAGAATGTGCTGACTGCAAAAACAAAAGGTCCAAGATGGGCTAGTGTAAAT 136
Db 61 CAGAGAACAGAGAATGTGCTGACTGCAAAAACAAAAGGTCCAAGATGGGCTAGTGTAAAT 120

Qy 137 TAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGAGTCTCGGGGTACACAT 196
Db 121 TAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGAGTCTCGGGGTACACAT 180

Qy 197 CGAAGTTGATCTGCCACTCTGACACATGGCTCCCGAGCAGGTTGCATTATACAGT 256
Db 181 CGAAGTTGATCTGCCACTCTGACACATGGCTCCCGAGCAGGTTGCATTATACAGT 240

Qy 257 CAATGGGAATGATAAAGCAATAGTTACTCGGAAGCAGAGCTACCCCAAACTATGATA 316
Db 241 CAATGGGAATGATAAAGCAATAGTTACTCGGAAGCAGAGCTACCCCAAACTATGATA 300

Qy 317 GAGTTGGAATTGAGAAATTTTATACGTGCAAAATGATGAAGAGAGAGATGGGTTCTAGAG 376
Db 301 GAGTTGGAATTGAGAAATTTTATACGTGCAAAATGATGAAGAGAGAGATGGGTTCTAGAG 360

Qy 377 GGGAAAGGCTAGATCACTCTAGCTGAGCAGGAGGACGGGAAATCTGTGGAGAGAA 436
Db 361 GGGAAAGGCTAGATCACTCTAGCTGAGCAGGAGGACGGGAAATCTGTGGAGAGAA 420

Qy 437 GTGGCCGGGATATGAGCATGAGCATAGTAGTCTGTAAATTTGTTGAGGAGAGA 496
Db 421 GTGGCCGGGATATGAGCATGAGCATAGTAGTCTGTAAATTTGTTGAGGAGAGA 480

Qy 497 AAATCTATTCAGCATCTAGAAACAAGAAATATGTTGTCGCAACGAGAAATATCTCCCG 556
Db 481 AAATCTATTCAGCATCTAGAAACAAGAAATATGTTGTCGCAACGAGAAATATCTCCCG 540

Qy 557 TGCCTCCCAAGGACCCAGTCCAGGTATTAAGCCACAGCAAAATGAGTCTGCAGCTA 616
Db 541 TGCCTCCCAAGGACCCAGTCCAGGTATTAAGCCACAGCAAAATGAGTCTGCAGCTA 600

Qy 617 CTCAGTAGAGAGGAGAGAAACAAGCAGTAATTTGTCACCCAGCATCAGATCTCCAAAG 676
Db 601 CTCAGTAGAGAGGAGAGAAACAAGCAGTAATTTGTCACCCAGCATCAGATCTCCAAAG 660

Qy 677 TGGATTTTGTCTACTGATCTGTTTAACTGATCAATGATGATGATGATGATGATGATGAT 736
Db 661 TGGATTTTGTCTACTGATCTGTTTAACTGATCAATGATGATGATGATGATGATGATGAT 720

Qy 737 CAGAGGCAACTCTCTGGCGATCTCTGCGGATGATTAATCATGGCTGGCTTTCAGTCTG 796
Db 721 CAGAGGCAACTCTCTGGCGATCTCTGCGGATGATTAATCATGGCTGGCTTTCAGTCTG 780

Qy 797 CTGGAAGTGGTCAAAACGGCAGAGAAAATTTGTACAGCCAAAGCTGCTGAGAGCAGTCTTC 856
Db 781 CTGGAAGTGGTCAAAACGGCAGAGAAAATTTGTACAGCCAAAGCTGCTGAGAGCAGTCTTC 840

Qy 857 CTCAGGCTTCAATCTTCTGACCTTTGAGGATTTGTTTAAAGCACACACCTAATTTTAACTC 916

| | | | |
|------------|--|--|------|
| Db | 841 | CTCCAGCTTCATCTTCTGAGATTTGTTTAAGGACACACCTAATTTAAACAATC | 900 |
| Qy | 917 | AAACAAGCACAAGAGTGAAGGAGATATCATGAGCTGTTTGAAGAGACGATATAG | 976 |
| Db | 901 | AAACAAGCACAAGAGTGAAGGAGATATCATGAGCTGTTTGAAGAGACGATATAG | 960 |
| Qy | 977 | TATCGCTTTTGCATGATCAGCAACAGTTCGTATGCTCGCTCAGCAGCAGCCCTTT | 1036 |
| Db | 961 | TATCGCTTTTGCATGATCAGCAACAGTTCGTATGCTCGCTCAGCAGCAGCCCTTT | 1020 |
| Qy | 1037 | ACATGCTGTCAGGCAAGAGTTCGTGAGGCACTCAACCGCGTGAATCAACAGCTATTG | 1096 |
| Db | 1021 | ACATGCTGTCAGGCAAGAGTTCGTGAGGCACTCAACCGCGTGAATCAACAGCTATTG | 1080 |
| Qy | 1097 | CTAATGCTTTTAAAGTTCGTGCAAAATTTGTTCAACCGCGGCTACAGATCCCG | 1156 |
| Db | 1081 | CTAATGCTTTTAAAGTTCGTGCAAAATTTGTTCAACCGCGGCTACAGATCCCG | 1140 |
| Qy | 1157 | GAATGCTTAAACCGGAGTTCGTGCAAAATTTGTTCAACCGCGGCTACAGATCCCG | 1216 |
| Db | 1141 | GAATGCTTAAACCGGAGTTCGTGCAAAATTTGTTCAACCGCGGCTACAGATCCCG | 1200 |
| Qy | 1217 | TGAACGCAACATGCAACAGAGACCCGCAACCGGCAAGAGACACTCTACATATCCCAT | 1276 |
| Db | 1201 | TGAACGCAACATGCAACAGAGACCCGCAACCGGCAAGAGACACTCTACATATCCCAT | 1260 |
| Qy | 1277 | CATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTATGACCCCAACTCAA | 1336 |
| Db | 1261 | CATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGTATGACCCCAACTCAA | 1320 |
| Qy | 1337 | CCGCTAAACCTCAGTCTATCGGCAACCCCAACCAACGACCAACCATCTTCAAT | 1396 |
| Db | 1321 | CCGCTAAACCTCAGTCTATCGGCAACCCCAACCAACGACCAACCATCTTCAAT | 1380 |
| Qy | 1397 | CAGGCAAGACTTTGATTTCTCTCTTCTGATGATGAATGTTTCAACAACATTTGA | 1452 |
| Db | 1381 | CAGGCAAGACTTTGATTTCTCTCTTCTGATGATGAATGTTTCAACAACATTTGA | 1436 |
| RESULT 6 | | | |
| LOCUS | AK067240 | 1451 bp mRNA linear PLN 24-JUL-2003 | |
| DEFINITION | Oryza sativa (japonica cultivar-group) cDNA clone:J013096N21, full insert sequence. | | |
| ACCESSION | AK067240.1 GI:32977258 | | |
| VERSION | AK067240.1 | | |
| KEYWORDS | FLI cDNA; CAP trapper. | | |
| SOURCE | Oryza sativa (japonica cultivar-group) | | |
| ORGANISM | Oryza sativa (japonica cultivar-group) | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. | | |
| AUTHORS | The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Sugiyama, A., Matsubara, K., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otono, Y., Tsunoda, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayashida, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oatato, N., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y. | | |
| TITLE | Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice | | |

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1451)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurokawa, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Takagi, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
PAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .1451
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013096N21"

FEATURES
source

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 1451;
Best Local Similarity 59.2%; Pred. No. 4.1e-74;
Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;

Qy

1 ATGACGAGAAAGCCCAACGCTCTCTTAGAGAGCTTAATGCCCGCATGAAAGATTCTTGAA 60

```

Db      278  ATGACGAGAGGCGCTCGCTCCACGAGGCTCAAGCCGACGACACAGAGATATTGGA 337
Qy      61  GGGCTTCTTAAACATCCAGAGAACAGAGATGTGCTGACTGCAAAACAAAAGTTCACAGA 120
Db      338  GGCCTCTCGGGCTGCTCGAATAGAGAAATGTGCAGCTGCAAGTCAAAAGGGTCTCTCGA 397
Qy      121  TGGGCTAGTGTAAATTTAGTGTCTTTATCTGCAATGCAAGTTCCTGGGATTCACAGAGT 180
Db      398  TGGCAAGTGTAAATCTTGGGATCTTTATATGATGCAAGTGTCTGGAATTCATAGAGC 457
Qy      181  CTGGGGTACACATATCGAAGGTTGCAATCTGCCACTCTGGACACATGGCTCCCCGAGCAG 240
Db      458  CTCGGGGTACATATCTAAGTAAGATCGCTACCTGATACATGCTGCCAGAGCAA 517
Qy      241  GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAGCAGAGCTTA 300
Db      518  GTTGCAATTTTCAATCAATGGGGAATGAAAGTCAATAGTACTCTGGGAAGCGGAGCTG 577
Qy      301  CCCCACAACTATGATAGTGTGGAATTTGCAATTTTATACGTGCAAGTATGAAGAGAG 360
Db      578  CCCCCTAATATGATAGGTTGGGATAGAGAACTTTCATCGTGCAAAATACGAGGACAG 637
Qy      361  AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTTCGAGCAGGAAACGGCGG 420
Db      638  AGATGGATACCGAGGATGGGTCTCATCAGACCATCATCTGCTGCTAGAGATGAGAGAGC 697
Qy      421  AATCTGTGGAGAGAGTGGGCGGGATATGAGATGGAATAGTAGTAGTCTCTGTAAT 480
Db      698  TCAGAGTCACAACTAGTGTAAACAGGGTGGCCATAATCAAAG-----ATCT 745
Qy      481  TTGTTTGGAGAGGAAAACCTATTCAGCATCTAGACAGAGAAATATGTTGCTGCAAG 540
Db      746  TCATTTGAGCAACACCGCATCTCACCGCTGCGAGTGAAGAAAATGCTCCAGTGGTTTCT 805
Qy      541  AGAATAAATCTTCCTCGTGCCTCCCGAGGACCCAGTCAGGTTTATAAGCCACACAGAAA 600
Db      806  AGGACGCCACCCAGGACCAACATCAGCCAAAGCAACACCATCAGTTCCTCAAGTTTCA 865
Qy      601  ATGGAGTCTGAGTACTCAGTAGAGAGAGGAGAGAAACAGCA---GTAATGTTGACCA 657
Db      866  CCTCCTCAACCCGAGAAATCACCTCTTAATGCAACACCCACCAAGGTTGAGAGGCCATCA 925
Qy      658  GCATCAGATCCTCAAGGTGGATTTTGTCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
Db      926  GTTGACCACTCTTAAGTGTGACTATGCAACTGATCTTCAACATGTTATCAATGGAT 985
Qy      718  GATTCCACTACAAATACCTCAGAGGCAACTCTCTGGCGATACCTCTCGCATGATAACTCA 777
Db      986  GGAACAACCGAGAAAGAGCAGAGTCA-----TCTTCTAATGATGATGTGCA 1033
Qy      778  TGGGCTGGCTTTCAGTCTGCGAAGTGGTCAACGCGCAGAGAAAATGTCACAGCCAG 837
Db      1034  TGGGAGGCTTCCAGTCTGAGAGCCAGTACCTAGCTCAGCAAAAAGGATTCCTGCTAAG 1093
Qy      838  CTGCTGAGAGCAGTCTCTCCAGCTTCACTCTCTGAC-----TTTGAGGATTTGTT 891
Db      1094  CCAGTAGAAGTAGGCCCTTACATCAGGCATAGAGCACTTGTTTAAAGATTCACCA 1153
Qy      892  AAGGACACACTAATTTAAACAATCAACAGCACCAAAAGATGTGAAAGCCGATATCATG 951
Db      1154  GCTGTGACATGTCCTCAGCTCCAGCTCCGCCACAGTAATGTGAAGATGATATCATG 1213
Qy      952  AGCCTGTTGAGAGAGCAATATAGTATCCCTTTTGGCCATGATCAGCAACAGGTTGCT 1011
Db      1214  AGTCTGTTTGAAGTCCAGTATGTTTGGCCATATGCTGTCCAGCAGCAGCAACTTGCT 1273
Qy      1012  ATGCTCGCTCAGAGCA 1028
Db      1274  TTTATGACCCCCAGCA 1290

```

. RESULT 7

AK067125
LOCUS
DEFINITION
ORYZA SATIVA (japonica cultivar-group) cDNA clone: J013098J09, full insert sequence.
AK067125
GI: 32977143
FLI CDNA; CAP trapper.
ORYZA SATIVA (japonica cultivar-group)
ORYZA SATIVA (japonica cultivar-group)
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

REFERENCE
AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team., Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuka, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Sciences 301 (5631), 376-379 (2003)
22752273
12869764

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 1741)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kishimoto, N., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Nami, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuka, K., Ohtsuka, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE
JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel.81-29-838-7007, Fax:81-29-838-7007)

COMMENT
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuka, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akai, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
1. .1741
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013098J09"

FEATURES source

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 1741;
Best Local Similarity 59.2%; Pred. No. 4.1e-74;
Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;

QY 1 ATGAACGAGAGAGCAACGCTCTTAAGAGCTTAATGCGCGCATAGAAAGATTCTGAA 60
DB 43 ATGAACGAGAGAGCTCGCTCTCAAGAGCTCAAGCCACGACACAGAGATATGGA 102
QY 61 GGGCTTCTTAAATCCAGAGACAGAGAAATGTGCTGATGTCGCAAAACAAAGGTCGA 120
DB 103 GGCCTCTCGCGCTGCTGAGAAATGAGAAATGTGCGAGCTGCAAGTCAAGGCTCTCGA 162
QY 121 TGGCTAGTGTAAATTTAGTATCTTTCTCGATGCAATGTTCTGGGATTCACAGGAGT 180
DB 163 TGGCAAGTGTAAATCTTGGGATCTTTATATGATGCAAGTGTCTGGAAATTCATAGAAGC 222
QY 181 CTCGGGTTACATATCGAAGTTCGATCTGCGACTCTGACACATGCTCCCGCAGCAG 240
DB 223 CTCGGGTTACATATCTAAGTAAGTTCGGTACCTGCTGATACATGCTGCCAGCA 282
QY 241 GTTGCAATTTATACGTAATGGAATGATAAGCAATAGTATCTGGAAGCAGAGCTA 300
DB 283 GTTGCAATTTATCAATCAATGGAATGAAAGTCAATAGCTACTGGAAGCGGAGCTG 342
QY 301 CCCCACAACTATGATAGTGGAAATGGAATTTATACGTGCAAGTATGAAGAGAG 360
DB 343 CCCCCTAATATGATGAGTGGGATGAGAACTTCTCCGTCGCAAAATACGAGACAG 402
QY 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACTCTCTAGAGTCGAGCAGGAACGGCG 420
DB 403 AGATGATACCGAGGAATGGGTATCAAGACCATCATCTGCTGCTAGAGTGAAGAGC 462
QY 421 AAATCTGTGGAGAGTGGCGGGATATAGCATGGAATGAGTATGATGCTCTTAAT 480
DB 463 TCAGAGTCACAACTAGTGTAAACAGGGGTGCCATATCAAG-----ATCT 510
QY 481 TGTGTTGAGGAGGAAACTATTCAGCATCTAGAACAGAAATATGTTGCTCAAG 540
DB 511 TCAATTTGAGCAACACCGCACTTCCAGCTGAGCAAAATTTGCTCCAGTGGTTCT 570
QY 541 AGAATAAATCTTCCGTGCTCCCAAGGACCGAGTCAGGTTATTAAGCCACAGCAGAA 600
DB 571 AGAGCCCAACCGACCAATCATGCGCAAAAGACACACATCATGTTCCAGGTTTCA 630
QY 601 ATGGAGTCTGCAAGTCTCTCAGTAGAGGGGAGAAACAGCA---GTAATTTGTCACCA 657

DB 631 CTCTCTCAACCCGAGAAATACCTCTTAATGAAACACACCAACCAAGGTTGAGAGGCCATCA 690
QY 658 GCATCAGATCTCTCAAGAGTGGATTTTCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
DB 691 GTTGCAACCACTCTTAAGTTGACTATGCACTGATCTTCAACATGTTATCAATGGAT 750
QY 718 GATTGCACTACAAATACCTCAGAGGCAACTCTCTGCGGATCTCTCTGCGGATGATCACTCA 777
DB 751 GGAACAACCGAGAAAGAGCAGAGTCA-----TCTTCTTAATGATGATAGTGCA 798
QY 778 TGGGCTGGCTTTTCTAGTCTGCTGGAGTGGTCAAAAGGCGAGAGAAATTTGTACAGCCCAAG 837
DB 799 TGGGAGGGCTTCCAGTCTGCGAGAGCCAGTACTAGCTCAGACAAAAGGATTTCTCTTAAG 858
QY 838 CTCTGCTGAGAGCAGTCTCTCTCCAGCTTCACTCTCTGAC-----TTTGAGGATTTGTT 891
DB 859 CCAGTAGAAGTAGCCCACTCTACATCAGGCATAGAGGACTTCTTTAAGATTTCACCA 918
QY 892 AAGGACACACCTAATTTAACTCAACAGCACCACAAAGATGTGAAGGGGATATCATG 951
DB 919 GCTGTGACAGTATCTCTAGCTCCAGCTGCCCCACAAAGTAAATGTGAAGATGATATCATG 978
QY 952 AGCTGTTTGAAGAGAGCAATATAGTATCGCTTTTGCATGTCATCAGCAACAGCTTGCT 1011
DB 979 AGTCTGTTGAGAAAGTCCAGTATGTTTGGCCATATGCTGCTCCAGCAGCAGCAACTTGCT 1038
QY 1012 ATGCTCGCTCAGCAGCA 1028
DB 1039 TTTATGACCCCGCAGCA 1055

AK068892 2157 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone: J023001G22, full insert sequence.

RESULT 8
AK068892
LOCUS AK068892.1 GI:32978917
DEFINITION FLI cDNA; CAP trapper.
ACCESSION AK068892
VERSION 1
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1
The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team.; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ontsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Teunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kuwumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN.; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

TITLE
japonica rice
Science 301 (5631), 376-379 (2003)
JOURNAL MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 2157)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsuura, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, K., Namiki, T., Narikawa, R., Nishikawa, J., Nishi, K., Oka, M., Ooka, H., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Sakai, C., Sakai, K., Saito, N., Ota, Y., Otsu, Y., Ry, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikawa, J., Oka, M., Ry, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagasaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, P., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

COMMENT

TITLE
JOURNALFEATURES
Source

1..2157
/organism="oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/clone="J023001G22"

ORIGIN

Query Match 21.1%; Score 307; DB 8; Length 2157;
Best Local Similarity 59.2%; Pred. No. 4.2e-74;
Matches 614; Conservative 0; Mismatches 390; Indels 33; Gaps 4;

QY 1 ATGACGAGAGAGCCATCTCTCTAGGAGCTTAATGCCCGCATAGAAAGATTCCTGAA 60
|||||
DB 277 ATGAACGAGAGAGCCCTCCCTCTCCAGGAGCTCAAGCCAGCAAGAGATATTGAA 336
|||||

QY 61 GGGCTCTTTAAACATCCAGAGACAGAGAAATGTGCTGACTGCAAAACAAAGGTCGAAGA 120
|||||
DB 337 GGCCTCTCGGCTGCCTGAGATAGAGAAATGTGCAGACTGCAAGTCAAGGGTCTCGA 396
|||||

QY 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
|||||

DB 397 TGGCAAGTGTAAATCTTTGGATCTTTATATGCAATGCAAGTGTCTGGAAATTCATAGAAC 456
QY 181 CTCGGGGTACACATATCAAGAGTTTCGATCTCCACTCTCGGACACATGCTCCCCGAGCAG 240
|||||
DB 457 CTCGGGGTACATATATCTTAAGTAAAGTCCGCTACCTTGGATACATGCTGCCAGGACAA 516
|||||

QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
|||||
DB 517 GTTGCAATTTATCAATCAATGGGAAATGAAAAGTCAAAATAGTACTCTGGGAAACGAGCTG 576
|||||

QY 301 CCCCCAACTATGATAGATTTGGAAATTTAGAGATTTTATACGTGCAAAAGTATGAGAGAG 360
|||||
DB 577 CCCCTTAACATATGATAGGTTGGGATAGAGAACTTCATCCGTGCAAAATAGCAGGACAG 636
|||||

QY 361 AGATGGGGTTCTAGAGGGGAAAAGGCTAGATCACCTCTAGAGTCGAGCAGAGAAACGGCGG 420
|||||
DB 637 AGATGGATACCGAGGAATGGGTCAACAAGCCATCATCTGGTGTAGAGATGAGAGAGC 696
|||||

QY 421 AAATCTGTGGAGAGAAGTGGGCCCGGATATAGAGCATAGGACATAGTAGTAGTCTCTGTAAT 480
|||||
DB 697 TCAGAGTCACAAACTAGTGTGTTAAACAGGGGTGGCCATAATCAAAG-----ATCT 744
|||||

QY 481 TTGTTTGGAGAGAGAAAACCTATTTCAGCATCTAGAACAGAAATAATGTTGCTGCAACG 540
|||||
DB 745 TCATTTGAGCAACACCGCACTTCACCAAGCTGCAAGTGAAGCAAAATTCCTCAGTGTGTTCT 804
|||||

QY 541 AGAATAAATCTTCCCGTGCCCTCCCAAGGACCCAGTCAAGTTATAAAGCCACAGCAGAAA 600
|||||
DB 805 AGGACGCCACCCAGGACCCACATCAGCCAAAGCACAACCATCAGTTCCTCCAGGTTTCA 864
|||||

QY 601 ATGAGTCTGCAGCTATCTCCAGTAGAGAGGAGGAGAAACAGCA---GTAAATTTGTCACCA 657
|||||
DB 865 CGTCTCAACCCGAGAAATCACCTCTCTAAATGCAACACCCAAAGGTTGAGAGGCCATCA 924
|||||

QY 658 GCATCAGATCTCCAAAGTGGATTTTGCTACTGATCTGTTTAAATCATGCTATCAATGAT 717
|||||
DB 925 GTTGACCACTCTTAAGTTGACTATGCAACTGATCTCTCAACATGTTATCAATGAT 984
|||||

QY 718 GATTTCGACTACAAATACCTCAGAGGCAACTCTCTGGCGATACCTCTGGCGATGATAACTCA 777
|||||
DB 985 GGAACAACCGAAGAAAGCAGAGTCA-----TCTTCTAATGATGATAGTAGTCA 1032
|||||

QY 778 TGGGCTGCTTTTCACTGCTGGAAGTGTCAAAAGGCGAGAGAAATTTGTACAGCCCAAG 837
|||||
DB 1033 TGGGAGGGCTTCCAGTCTGACAGCCAGTACCTAGCTCAGACAAAAGGATTTCTGCTAAG 1092
|||||

QY 838 CCTCTCAGAGCAGTCTCTCTCCAGCTTCATCTTCTGAC-----TTTGAGGATTTGTT 891
|||||

DB 1093 CCAATGAGAAAGTAAGCCCGAGTCTACATCAGGATAGAGGACTTGTTTAAAGATTCACCA 1152
|||||

QY 892 AAGGACACACCTAATTTAAACAACCTCAACAAGCACCAAAAGATGTAAAGGCGCATATCATG 951
|||||

DB 1153 GCTGTGACATGATCTCTCAGTCCAGCTGCCCGCCACAAGTAAATGTGAAGATGATATCATG 1212
|||||

QY 952 AGCTGTTTGAAGAGAGCAATATAGTATCGCTTTTGGCATGCAATCAGCAACAGTTGCT 1011
|||||

DB 1213 AGTCTGTTGAGAAAGTCCAGTATGCTTTCGCCATATGCTGCTCCAGCAGCAGCAACTTCT 1272
|||||

QY 1012 ATGCTCGCTCAGCAGCA 1028
|||||

DB 1273 TTTATGACCCCGCCAGCA 1289
|||||

RESULT 9
LOCUS AB010695/1 78596 bp DNA linear PLN 14-FEB-2004
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone MDK4.
ACCESSION AB010695
VERSION AB010695.1 GI:2828180
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

ACWLLGVQVRSKCLKEOCENTIGCDLRLMLSCKEPVVYGVTTVMVLDRLARLAWAQHQA
 RVNIDINTNTYQWYQITQVSESREKILPIFWMLMTSTLTFGNLSTSTWSEV
 VFTIIVLTSGLLVTMLIGNITKVLHATTSKKQAMHLKMRNIEMWKKHDLPLTFQROR
 VRYRQSWAAMRGDEQVQMLPEGLRRDKYHLCIDLVRQVPLFQHMDDLVLLEN
 CDVKYSLTFKTEGITEQEGDAQRMFLFVRGHLOSSLDRGVKSCCMGLPGNPSGDB
 LLSKRPFPFVERLPPSSSTLVTLETTAFLGDAEDVKYVQYHFRYTVNEKVKRSAR
 YVSGWRTPAAVAQVLAWRRYKHLRLTSLTSFIRPRPLSRCASLGEDKRLYLAAIIT
 SPKNPDDDDDI
 join(35464,.35471,35682,.35817,36015,.36172,36315,.36405,
 36638,.36647,36728,.36796,36865,.36956,37045,.37147,
 37247,.37354,37476,.37538,37610,.37681,37785,.37862,
 38060,.38122,38211,.38271,38753,.38835,38979,.39092,
 39212,.39289,39376,.39447,39534,.39766,39879,.40002,
 40161,.40364,40479,.40511)
 /note="emb|CAB50793.1
 gene id:MDK4.8"
 /codon start=1
 /evidence=not experimental
 /product="DNA repair and meiosis protein Mrel1"
 /protein_id="BAB10749.1"
 /db_xref="GI:9759499"
 /translation="MRSRDPDTRLVLTVDCHLGYMEKDIRRHDSFKAFEEICISIA
 BEKVQDFPGLGDFPHENKGRPTTLVKAIEILRRHLNDKPVQVQSDOTVNFQNAF
 GQVAYDPHFNGVLPVSIHGNHDDPAGVNLSDAILDLCNLYVYGVKLGSGVG
 QITLYPLMKGSTVALYGLGNIRDERLNRMFQTPHAYQMRREVEGGCDVDWFNI
 LVLHNRVKNPNKPAISHFLPFRFLDFIVGHEHECLIDPOEVSGMGFHTIPQSSVA
 TSLIDGSGKPHLLEIKNQYRPTKPIPLTSVRPFYTEIVLKDSESDIDPNQNSIL
 EHLCDVNRNLEKAKAVNSEIKLPLVRIKDYSGFMTPINQRFQVKYKVPANP
 DILAFPSKAKGSEANDDSERLPEELNQONTALVAESNLKWEILPVNLDVALH
 NFNVKDDKLAFLYSCOVYNLQTRGKLAKSDPAKPFEDDLILKYGECLERLRKDSRTR
 PTGSGQLSTGLSENLYKSGSGIANASFSDEDDTQMSGLAPPTRGRSGSNTTR
 GRAKAPTRGRGKASAMKGTITLDSLGFQSQRSASAAASAASTIGEDDDVS
 PSSVEEPDEFNKPDSSEDDSTKGKGRKPAITKRGGRSGTSKRGKRNISSSL
 NRLLSDDDEDEDEKLNKSKQPRVTRNYGNR"
 join(41186,.41311,41410,.41571,41636,.42145)
 /note="gb|AAD28773.1
 gene id:MDK4.9"
 /codon start=1
 /evidence=not experimental
 /product="LhcB3 chlorophyll a/b binding protein"
 /protein_id="BAB10750.1"
 /db_xref="GI:9759500"
 /translation="MSTPTSSSVLTPTTFLGQTKASSFNPLRDVVVSLGSPKVTMGN
 DLWTGPRRVKLYGFPFVSQTFSLTGEPDGYGMDTAGLSADPEAFKAKRALEVIHGRW
 AMLGAFGCIITPEVLQKVRVDFKEPVNFKAGSQIFSEGGDLPLGNPLVHAQSILAVL
 GFQVILMGLVEGFRINGLDGVGEGNDLYPGQGYFDPLGLADDPVTPFAELKVKYIKNGR

| Query Match | 17.6%; Score 255.4; DB 8; Length 78596; |
|---------------------------|---|
| Best Local Similarity | 60.1%; Pred. No. 1.2e-59; |
| Matches 877; Conservative | 0; Mismatches 1; Indels 581; Gaps 4; |
| Qy 575 | GTCAGGTTATAAGCCACAGCAGAAAATGGAGTCTGCAGCTACTCCAGTACAGAGGAGGA 634 |
| Db 61482 | GACAGGTTATAAGCCACAGCAGAAAATGGAGTCTGCAGCTACTCCAGTACAGAGGAGGA 61423 |
| Qy 635 | AACAGCAGTAATAATGTTCCACAGCATCAGATCCTCCAAAGGTGGATTTTCTACTGATC 694 |
| Db 61422 | AACAGCAGTAATAATGTTCCACAGCATCAGATCCTCCAAAGGTGGATTTTCTACTGATC 61363 |
| Qy 695 | TGTTTAAATGCTATCAATGGANGATTGCATCAAAATACCTCAGAGGCAACTCTCTGGCG 754 |
| Db 61362 | TGTTTAAATGCTATCAATGGANGATTGCATCAAAATACCTCAGAGGCAACTCTCTGGCG 61303 |
| Qy 755 | ATACCTCTGCCGATCAATCAATCGGCTGGCTTCAGT----- 793 |
| Db 61302 | ATACCTCTGCCGATCAATCAATCGGCTGGCTTCAGTGTATGTATCTATCTAGCATGT 61243 |
| Qy 794 | ----- 793 |
| Db 61242 | GATTAATCTAATCGTGCCACATTTTTTCTATCTTGTAACTTTATGTCAATTTGTTCTT 61183 |
| Qy 794 | -----CTGCTGG 800 |

| | | |
|-----------|---|--|
| REFERENCE | 3 (bases 1 to 3504) | |
| AUTHORS | Klein, S. and Gerhard, D.S. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (19-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA | |
| REMARK | NIH-MGC Project | |
| COMMENT | Contact: XGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca | |
| FEATURES | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 155 Row: 9 Column: 13 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein. Location/Qualifiers 1..3504 /organism="Xenopus laevis" /mol_type="mRNA" /db_xref="taxon:8355" /clone="MGC:80897 IMAGE:5515434" /tissue_type="Embryo, stage 31/32, Xenopus" /clone_lib="NICHD_XGC_Emb4" /lab_host="DH108" /note="Vector: pCMV-SPORT6" 1..3504 /gene="MGC80897" /db_xref="LocusID:446881" 384..1436 /gene="MGC80897" /codon_start=1 /product="MGC80897 protein" /protein_id="AAH77937.1" /db_xref="GI:50417734" /db_xref="LocusID:446881" /translation="MTTSRERKAQKLNHQHQAIIISRLREEDNKYCADCEAKGPRWA SNLGVFCIRGAGTHRLNLGVHISRVKSNLDQWTPQIQCMQDMGNTKARIYEANL PENFRPOTDSVEFFIRDKYRKYYKDNATVSGTEKQKKERRKARPEKRWKVP VGADVLKKEPLVSKAPSKSPETLVLLGLDAPALNATSAFVAAVNDLIDIF GPMVNPPIVPSAAPTQVTSNPASSVSSSSGLDLFTDDTTKTDTPKPLSKDSI LSLYTANQQQNTTGMFGSSPFMPSYQSPASFCAPFPMGVPIAAGGIMGMNGPSV PMGTGECDDGRNGYA" | |
| ORIGIN | Query Match 7.7%; Score 111.2; DB 5; Length 3504; Best Local Similarity 59.5%; Pred. No. 1.7e-19; Matches 188; Conservative 0; Mismatches 128; Indels 0; Gaps 0; | |
| QY | 23 CTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAAGGGTCTTTAAACATCCAGAGA 82 | |
| Db | | |
| QY | 412 CTCAGAAGCTGAATGAGCAGCATCAGGCCATCTTGTCCCGATGCTTCGGGAGGAAGACA 471 | |
| Db | | |
| QY | 83 ACAGAGATGTCTGACTGCAAAACAAAGGTCGAGATGGCTAGTGTAAATTATTTAGTA 142 | |
| Db | | |
| QY | 472 ACAAGATTGTCTGACTGTAAGCAAAAGGTCCTAGGTGGGCTTCTTGGAACCTAGGAG 531 | |
| Db | | |
| QY | 143 TCTTTATCTGCAATGTTCTGGGATTTCACAGGAGTCTCGGGGTGACATATATCGAAGG 202 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAGGATATGG 651 | |
| Db | | |
| QY | 263 GAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTG 322 | |
| Db | | |
| QY | 652 GAAATCAAGGCAAGAGCTATATACGAGTATCTGCCAGAAAACCTCCGAGAGCCTC 711 | |
| Db | | |
| QY | 323 GAAATGAGAAATTTTAT 338 | |
| Db | | |
| QY | 712 AGACTGATCAATCTGT 727 | |
| Db | | |
| QY | 532 TTTTATGTCATTGGTTGTGCTGGAATTCATAGAAACCTTGGAGTTCATATATCAAGAG 591 | |
| Db | | |
| QY | 203 TTCGATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCATTTATACAGTCAATGG 262 | |
| Db | | |
| QY | 592 TGAATCTGTTAAATTTGGATCAATGGACTCCAGAACAAATTCAGTGCATGAG | |


```

/genes="SMAP1"
/notes="synonyms: FLJ13159, SMAP-1"
/db_xref="taxon:9606"
116...>554
/genes="SMAP1"
/codon_start=1
/product="SMAP1 protein"
/protein_id="AAH08672.1"
/db_xref="GI:34189699"
/db_xref="LocusID:60682"
/translation="WATRSCKRAQKLNQHOHLILSKLLREDNKKYCADCEAKGPRWA
SWNGVFCIRCAIGHNRLGVHISRVKSVNLDQWTAEQIQCMQDMGNTKARLLYEANL
PENFRPQTDQAVEFFIRDKYKKYDKNAITNKEKEKKREKPEKPAKP
LTAFLKQKQQLPEPKKSTSPKAAEPTVLLGLDGPVAVPTVNGTNTVPLNDLDDI
FGPMISNPLATVMPPAQGTSPAPAAATLSTVTSQDLDLFTQETTKSEVAKQLSKD
SILSLYGTGTIOQSTPGVFMGPTNIPTSOAPAAFOGFPSPMGVPAAPGLIGNVMG
QSPSMVGMPPNPGFMGNACTGUMPLQNVVVGPOGVMGQMGAPQSKFLPQAQQPQW
SLSQMNQMGAMSSISATPTAGFGQSPSTTAGWSGSSGQTLSTQLMK"
misc_feature
149..421
/genes="SMAP1"
/notes="COG5347; Region: GTPase-activating protein that
regulates ARFs (ADP-ribosylation factors), involved in
ARF-mediated vesicular transport [intracellular
trafficking and secretion]"
/db_xref="CDD:COG5347"

ORIGIN
Query Match 7.4%; Score 107.2; DB 9; Length 554;
Best Local Similarity 60.1%; Pred. No. 2.1e-18;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 23 CTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAAGGGCTTCTTAAACATCCAGAGA 82
DB 144 CTCAGAGCTGAACGAGGAGCACCGCTCATCTCTCCAGCTTCTGAGGGAGGAGACA 203
QY 83 ACAGAGAATGTGCTGATCTGCAACAAACAAAGGTCCTCAAGATGGGCTAGTGTATTTAGGTA 142
DB 204 ACAAGTACTGCCGCACTCGAGGCGCAAGGTCCTCGATGGGCTTCTTGAATATTTGGTG 263
QY 143 TCTTTATCTGATGCATGATTTCTGGATTCACAGAGTCTCGGGTACACATATCGAAG 202
DB 264 TGTTTATTTGCATCAGATGCTGGAATTCATAGAAATCTTGGGGTTCATATATCCAGG 323
QY 203 TTGCATCTGCCACTCTGGACACATGGCTCCCGAGCAGGTTCCTTATATACAGTCAATGG 262
DB 324 TCAATCAGTCAACCTAGACCATGGACAGCAGACAGATACAGTGCATGCAAGATATGG 383
QY 263 GAAATGATAAACAATAGTTACTGGGAGCAGAGCTACCCCAAACTATGATAGA 318
DB 384 GAAATACTAAAGCAAGACTACTCTATGAAGCAATCTCCAGAGAACCTTTCGAAGA 439

RESULT 15
AY055004
LOCUS
DEFINITION
Homo sapiens stromal membrane-associated protein SMAP1B (SMAP1)
mRNA, complete cds, alternatively spliced.
ACCESSION
AY055004
VERSION
AY055004.1 GI:16303737
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2223)
Marcos,I., Borrego,S., Rodriguez de Cordoba,S., Galan,J.J. and
Antinolo,G.
Cloning, characterization and chromosome mapping of the human SMAP1
gene
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2223)
Marcos,I., Borrego,S., Rodriguez de Cordoba,S., Galan,J.J. and
Antinolo,G.
TITLE
Direct Submission
JOURNAL
Submitted (28-AUG-2001) Unidad de Genetica Medica y Diagnostico
Prenatal, Hospitales Universitarios Virgen del Rocío, Avda. Manuel
Siurot s/n, Sevilla 41013, Spain
FEATURES

```

```

1..2223
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
1..2223
/genes="SMAP1"
/genes="SMAP1"
/genes="SMAP1"
/genes="SMAP1"
/translation="WATRSCKRAQKLNQHOHLILSKLLREDNKKYCADCEAKGPRWA
SWNGVFCIRCAIGHNRLGVHISRVKSVNLDQWTAEQIQCMQDMGNTKARLLYEANL
PENFRPQTDQAVEFFIRDKYKKYDKNAITNKEKEKKREKPEKPAKP
LTAFLKQKQQLPEPKKSTSPKAAEPTVLLGLDGPVAVPTVNGTNTVPLNDLDDI
FGPMISNPLATVMPPAQGTSPAPAAATLSTVTSQDLDLFTQETTKSEVAKQLSKD
SILSLYGTGTIOQSTPGVFMGPTNIPTSOAPAAFOGFPSPMGVPAAPGLIGNVMG
QSPSMVGMPPNPGFMGNACTGUMPLQNVVVGPOGVMGQMGAPQSKFLPQAQQPQW
SLSQMNQMGAMSSISATPTAGFGQSPSTTAGWSGSSGQTLSTQLMK"
3'UTR
1392..2223
/polya_signal
/genes="SMAP1"
/genes="SMAP1"

ORIGIN
Query Match 7.4%; Score 107.2; DB 9; Length 2223;
Best Local Similarity 60.1%; Pred. No. 2.2e-18;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 23 CTAAGGAGCTTAATGCCGCCCATAGAAAGATTCTTGAAGGGCTTCTTAAACATCCAGAGA 82
DB 97 CTCAGAGCTGAACGAGGAGCACCGCTCATCTCTCCAGCTTCTGAGGGAGGAGACA 156
QY 83 ACAGAGAATGTGCTGACTGCAAAAACAAAAGGTCCAAGATGGGCTAGTGTTTAATTTAGGTA 142
DB 157 ACAAGTACTGCCGCACTCGAGGCGCAAGGTCCTCGATGGGCTTCTTGAATATTTGGTG 216
QY 143 TCTTTATCTGATGCATGATTTCTGGATTCACAGAGTCTCGGGTACACATATCGAAG 202
DB 217 TGTTTATTTGCATCAGATGCTGGAATTCATAGAAATCTTGGGGTTCATATATCCAGG 276
QY 203 TTGCATCTGCCACTCTGCACACATGGCTCCCGAGCAGGTTCCTTATATACAGTCAATGG 262
DB 277 TCAATCAGTCAACCTAGACCATGGACAGCAGACAGATACAGTGCATGCAAGATATGG 336
QY 263 GAAATGATAAACAATAGTTACTGGGAGCAGAGCTACCCCAAACTATGATAGA 318
DB 337 GAAATACTAAAGCAAGACTACTCTATGAAGCAATCTTCCAGAGAACCTTTCGAAGA 392

Search completed: September 27, 2005, 05:03:34
Job time : 6240 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 02:51:39 ; Search time 282 Seconds
(without alignments)
8425.084 Million cell updates/sec

Title: US-10-630-518-1
Perfect score: 1452
Sequence: 1 atgaacgagaaagccaaagt.....gaatgttcacaaacattga 1452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCFUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 199.6 | 13.7 | 270 | 4 | US-09-313-294A-574 |
| 2 | 182.4 | 12.6 | 408 | 4 | US-09-640-211A-1383 |
| 3 | 107.2 | 7.4 | 2885 | 4 | US-09-774-528-3 |
| 4 | 94.2 | 6.5 | 1447 | 3 | US-09-484-970B-121 |
| 5 | 94.2 | 6.5 | 1976 | 4 | US-09-620-312D-515 |
| 6 | 86.2 | 5.9 | 309 | 4 | US-09-513-999C-2719 |
| 7 | 86.2 | 5.9 | 459 | 4 | US-09-621-976-845 |
| 8 | 82.6 | 5.7 | 1327 | 4 | US-09-620-312D-1104 |
| 9 | 78.2 | 5.4 | 1261 | 4 | US-09-855-323-4 |
| 10 | 74.8 | 5.2 | 514 | 4 | US-09-854-133-660 |
| 11 | 72.2 | 5.0 | 435 | 4 | US-09-640-211A-1584 |
| 12 | 68.6 | 4.7 | 6990 | 4 | US-09-620-312D-135 |
| 13 | 68.4 | 4.7 | 332 | 4 | US-09-513-999C-384 |
| 14 | 65.4 | 4.5 | 2119 | 4 | US-09-620-312D-35 |
| 15 | 63.4 | 4.4 | 7218 | 1 | US-08-232-463-14 |
| 16 | 62.8 | 4.3 | 333 | 4 | US-09-640-211A-1594 |
| 17 | 62 | 4.3 | 1455 | 4 | US-09-248-796A-4726 |
| 18 | 55.6 | 3.8 | 286 | 4 | US-09-313-294A-4791 |
| 19 | 55.6 | 3.8 | 543 | 4 | US-09-248-796A-4610 |
| 20 | 49.4 | 3.4 | 1257 | 4 | US-09-270-767-10482 |
| 21 | 47.8 | 3.3 | 505 | 4 | US-09-513-999C-1131 |
| 22 | 47.8 | 3.3 | 1254 | 4 | US-09-620-312D-804 |
| 23 | 47.4 | 3.3 | 213 | 4 | US-09-513-999C-22310 |
| 24 | 46.6 | 3.2 | 502 | 4 | US-09-828-303-3 |
| 25 | 46.6 | 3.2 | 2041 | 4 | US-09-828-303-11 |
| 26 | 45.4 | 3.1 | 1141 | 4 | US-09-806-708B-22 |
| 27 | 45.2 | 3.1 | 2712 | 4 | US-09-023-905A-11 |

| | | | | | | |
|----|------|-----|---------|---|---------------------|-------------------|
| 28 | 45.2 | 3.1 | 4595 | 4 | US-09-023-905A-9 | Sequence 9, Appli |
| 29 | 44 | 3.0 | 5330 | 4 | US-09-023-905A-1 | Sequence 1, Appli |
| 30 | 42.4 | 2.9 | 473 | 4 | US-09-471-276-723 | Sequence 723, App |
| 31 | 42.4 | 2.9 | 562 | 4 | US-09-621-976-338 | Sequence 338, App |
| 32 | 41.8 | 2.9 | 2949 | 4 | US-09-023-905A-8 | Sequence 8, Appli |
| 33 | 41.8 | 2.9 | 5954 | 4 | US-09-023-905A-6 | Sequence 6, Appli |
| 34 | 41.4 | 2.8 | 699 | 4 | US-09-248-796A-5451 | Sequence 5451, Ap |
| 35 | 41.2 | 2.8 | 3456 | 4 | US-09-023-905A-5 | Sequence 5, Appli |
| 36 | 41.2 | 2.8 | 4382 | 4 | US-09-023-905A-3 | Sequence 3, Appli |
| 37 | 40.4 | 2.8 | 235064 | 4 | US-09-949-016-15390 | Sequence 15390, A |
| 38 | 38.8 | 2.7 | 957 | 3 | US-09-134-001C-1852 | Sequence 1852, Ap |
| 39 | 38.8 | 2.7 | 1194 | 4 | US-08-710-279-1243 | Sequence 1243, Ap |
| 40 | 38.8 | 2.7 | 2998 | 4 | US-09-710-279-3506 | Sequence 3506, Ap |
| 41 | 38.8 | 2.7 | 44342 | 4 | US-09-949-016-12661 | Sequence 12661, A |
| 42 | 38.6 | 2.7 | 333 | 4 | US-09-640-211A-1761 | Sequence 1761, Ap |
| 43 | 38.6 | 2.7 | 1141 | 4 | US-09-806-708B-22 | Sequence 22, Appl |
| 44 | 37.2 | 2.6 | 832 | 4 | US-09-621-976-2813 | Sequence 2813, Ap |
| 45 | 36.8 | 2.5 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-09-313-294A-574
; Sequence 574, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 574
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549464HI
US-09-313-294A-574

| | | | | | | | |
|-----------------------|-------|---|---------|------------|----|--------|-----|
| Query Match | 13.7% | Score | 199.6 | DB | 4 | Length | 270 |
| Best Local Similarity | 83.7% | Pred. No. | 4.6e-57 | | | | |
| Matches | 226 | Conservative | 0 | Mismatches | 44 | Indels | 0 |
| Gaps | 0 | | | | | | |
| Qy | 92 | GTGCTGACTGCAAAACAAAAGTCCAGATGGCTAGTGTAAATTTAGGTATCTTTATCT | 151 | | | | |
| Db | 1 | GGCGAGCTGCAAGTCAAGGGTCTCGATGGCAAGTGTGAATCTAGTATCTTTATAT | 60 | | | | |
| Qy | 152 | GCATGCAATGTTCTGGGATTCACAGAGTCTCGGGGTACATATCGAAGGTTCAATCTG | 211 | | | | |
| Db | 61 | GCATGCAATGTTCTGGGATTCACAGAGTCTCGGGGTACATATCGAAGGTTCAATCTG | 120 | | | | |
| Qy | 212 | CCACTCTGCACATGCTCCCGAGCAGGTTCATTTATACAGTCAATGGGAAATGATA | 271 | | | | |
| Db | 121 | CCACCTCGATACATGCTGCGCAGAGCAAGTTCATTTATTCATCAATGGGAAACAAA | 180 | | | | |
| Qy | 272 | AAGCAAAATGTTACTGCGGAAGCAGAGCTACCCCAAACTATGATAGAGTTGGAATGAGA | 331 | | | | |
| Db | 181 | AAGCAAAATGTTACTGCGGAAGCAGAGCTACCCCAAACTATGATAGAGTTGGAATGAGA | 240 | | | | |
| Qy | 332 | ATTTTATACGTGCAAAAGTATGAAGAGAAGA | 361 | | | | |
| Db | 241 | ATTTTATACGTGCAAAAGTATGAAGAGAAGA | 270 | | | | |

RESULT 2
US-09-640-211A-1383


```
; Sequence 1383, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1383
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-640-211A-1383

Query Match      12.6%; Score 182.4; DB 4; Length 408;
Best Local Similarity 85.0%; Pred. No. 4.4e-51;
Matches 204; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 ATGACGAGAGAGCCAAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAGAAAGATTCTTGAA 60
   |||||
Db 169 ATGACGGGAGGCGCAAGCTCTCCAGGAGCTCAATGCCACCAAGAAAGATTCTCGNA 228
   |||||

QY 61 GGGCTTCTTAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGGTCCAAGA 120
   |||||
Db 229 GGGCTTCTCAAAATGCTGAGAACAGGGAGTGTGCTGATTGTAAGGCCAAAGGTCCAAGA 288
   |||||

QY 121 TGGGCTAGTGTAAATTTAGGTATCTTATCTGATCAATGTTCTGGGATTCACAGGAGT 180
   |||||
Db 289 TGGGCTAGTGTCAATTTAGGATATTTATGATGCAATGTTCAAGGAATCATAGAAGT 348
   |||||

QY 181 CTCGGGGTACACATATCGAAGGTTGCTGCTGCCACTCTGGACACATGCTCCCGGAGCAG 240
   |||||
Db 349 CTTGGGGTACACATATCGAAGGTTGCTGCTGCCACTCTGGACACATGCTCCCGGAGCAG 408
   |||||

RESULT 3
US-09-774-528-3
; Sequence 3, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 3
; LENGTH: 2885
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (120)...(1442)
US-09-774-528-3

Query Match      7.4%; Score 107.2; DB 4; Length 2885;
Best Local Similarity 60.1%; Pred. No. 6.7e-25;
Matches 178; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 23 CTAGGAGCTTAATGCCCGCCATAGAGAAATTTCTTGAAGGGCTTCTTAAACATCCAGAGA 82
   |||||
Db 148 CTCAGAGCTGACGAGCAGCAGCAGCTCATCTTATCCAAAGCTTCTGAGGGAGGAGACA 207
   |||||

QY 83 ACAGAGAAATGTGCTGACTGCAAAACAAAAGTCCAAAGTGGGCTAGTGTATTTAGTGA 142
   |||||
Db 208 ACAAGTACTGCGCGGACTGCGAGGCCAAAGGCTCTCGATGGGCTTCTCTGGAATATTGGTG 267
   |||||

QY 143 TCTTTATCTGCATGCAATGTTCTGGGATTTACAGAGAGTCTCGGGGTACACATATCGAAGG 202
   |||||
Db 268 TGTTTATTTGCATCAGATGTGCTGGAATTTATAGAAATCTTGGGGTTTCATATATCCAGGG 327
   |||||

QY 203 TTCGATCTGCCACTCTGGACACATGCGCTCCCGAGCAGGTTGCATTTATACAGTCAATGG 262
   |||||
Db 328 TCAATCAGTCAACTAGACCAGTGGACAGCAGACAGATACAGTGCATGCAAGATATGG 387
   |||||

QY 263 GAAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTACCCCCAAACTATGATAGA 318
   |||||
Db 388 GAAATACTAAAGCAAGACTACTCTATGAAGCCAACTTTCAGAGAACTTTCGAAGA 443
   |||||

RESULT 4
US-09-484-970B-121
; Sequence 121, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 121
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 227709 (1221361CBI)
US-09-484-970B-121

Query Match      6.5%; Score 94.2; DB 3; Length 1447;
Best Local Similarity 61.0%; Pred. No. 1.1e-20;
Matches 153; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 79 GAGAACAGAGAAATGTGCTGACTGCAAAACAAAAGSTCCAAAGTGGCTAGTGTATTAATTA 138
   |||||
Db 110 GATAACAAGTTTGTGCGGATTTGCCAGTCTAAAGGCCGCGATGGGCTCTTGGAAACATT 169
   |||||

QY 139 GGTATCTTTATCTGCATGCAATGTTCTGGGATTTACAGGAGTCTCGGGGTACACATATCG 198
   |||||
Db 170 GGTGTGTTTCATCTGCAATTCGATGTGCTGGAATCCACAGGAATCTGGGGGTGCACATATCC 229
   |||||

QY 199 AAGGTTTCGATCTGCCACTCTGGACACATGCGCTCCCGAGCAGGTTGCATTTATACAGTCA 258
   |||||
Db 230 AGGGTAAAGTCAAGTTAACTTCGACAGTGGACTCAAGAACAGATTCAGTGCATGCAAGAG 289
   |||||

QY 259 ATGGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTACCCCCAAACTATGATAGA 318
   |||||
Db 290 ATGGGAATGGAAGGCAACCGACTTTATGAAGCTATCTTCTCTGAGACCTTTTCGGGGA 349
   |||||

QY 319 GTTGAATTTGA 329
   |||||
```

Db 350 CCTCAGATAGA 360

RESULT 5

US-09-620-312D-515
Sequence 515, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 515

LENGTH: 1976

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (47)..(1336)

US-09-620-312D-515

Query Match 6.5%; Score 94.2; DB 4; Length 1976;

Best Local Similarity 61.0%; Pred. No. 1.3e-20;

Matches 153; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 79 GAGAACAGAGATGCTGACTGCAAAACAAAGGTCACAGATGGGCTAGTGTAAATTA 138

Db 116 GATAACAAAGTTTGTGCGAGATTGCCAGTCTAAAGGCGCGGATGGGCTCTTGGAAACAT 175

Qy 139 GGTATCTTTATCTGATCAATGTTCTGGATTCTCAGAGTCTCGGGGTACACATATCG 198

Db 176 GGTGTGTTTCATCTGATTCGATGCTGAATCCACAGGATCTGGGGGTGCACATATCC 235

Qy 199 AAGGTTTCGATCTGCCACTCTGGACACATGGCTCCCGGAGCAGGTTGCATTTATACAGTCA 258

Db 236 AGGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 295

Qy 259 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 318

Db 296 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 355

Qy 319 GTTGGAAATGA 329

Db 356 CCTCAGATAGA 366

RESULT 6

US-09-513-999C-2719

Sequence 2719, Application US/09513999C

Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59, US2, REG

CURRENT APPLICATION NUMBER: US/09/513,999C

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO 2719

LENGTH: 309

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 67..309

US-09-513-999C-2719

Query Match 5.9%; Score 86.2; DB 4; Length 309;

Best Local Similarity 59.0%; Pred. No. 1.8e-18;

Matches 148; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 79 GAGAACAGAGATGCTGACTGCAAAACAAAGGTCACAGATGGGCTAGTGTAAATTA 138

Db 46 GCGGAAGGAAACTGCTTAAATGATTTTAAAGGGCGCGATGGGCTCTTGGAAACAT 105

Qy 139 GGTATCTTTATCTGATCAATGTTCTGGGATTCTCAGGAGTCTCGGGGTACACATATCG 198

Db 106 GGTGTGTTTCATCTGATTCGATGCTGGAATCCACAGGATCTGGGGGTGCACATATCC 165

Qy 199 AAGGTTTCGATCTGCCACTCTTGACACATGGCTCCCGGAGCAGGTTGCATTTATACAGTCA 258

Db 166 AGGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 225

Qy 259 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 318

Db 226 ATGGGAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAACTATCATAGA 285

Qy 319 GTTGGAAATGA 329

Db 286 CCTCAGATAGA 296

RESULT 7

US-09-621-976-845

Sequence 845, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSER.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 845

LENGTH: 459

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 52..459

US-09-621-976-845

Query Match

5.9%; Score 86.2; DB 4; Length 459;

Best Local Similarity 59.0%; Pred. No. 2.4e-18;

| | Matches | 148; | Conservative | 0; | Mismatches | 103; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|------|--------|----|------|----|
| Qy | 79 | GAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCCAAAGATCGGCTAGTGTTAATTTA | 138 | | | | | | | |
| Db | 31 | GCGAAAGGAAAACTGCTTAAATGATTTTAAAGGGCCGCGATGGGCTCTTGGAAACATT | 90 | | | | | | | |
| Qy | 139 | GGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGTCTCGGGGTACACATATCG | 198 | | | | | | | |
| Db | 91 | GGTGTGTTTCATCTGCATTCGATGTCTGGATCCACAGGATCTCGGGGTGCATATCC | 150 | | | | | | | |
| Qy | 199 | AAGTTTCGATCTGCCACTCTTGGAACACATGTGCTCCCGAGCAGGTTGCAATTAACAGTCA | 258 | | | | | | | |
| Db | 151 | AGGGTAAAGTCAGTTAACTCTCGACCAGTGGACTCAAGAAACAGATTCAGTGCATGCAAGAG | 210 | | | | | | | |
| Qy | 259 | ATGGGAATGATAAGCAAAATAGTTTACTTGGGAAGCAGAGCTACCCCAACATATGATAGA | 318 | | | | | | | |
| Db | 211 | ATGGGAATTGGAAAGGCNAACCGACTTTATGNAAGCCTATCTTCTGAGACCTTTTCGGGGA | 270 | | | | | | | |
| Qy | 319 | GTTCGAATTTGA | 329 | | | | | | | |
| Db | 271 | CCTCAGATAGA | 281 | | | | | | | |

```

RESULT 8
US-09-620-312D-1104
; Sequence 1104, Application US/09620312D
; Patent NO. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1104
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (251)..(1300)
US-09-620-312D-1104

```

| | | | |
|----|-----|---|-----|
| Qy | 165 | TGGGATTACAGAGGAGTCTCGGGGTGACATATCGAAGTTTCGATCTGCCACTCTCGACAC | 225 |
| Db | 166 | TGGAAATCCACAGGAATCTGGGGGTGCATATCCAGGGTAAAGTCAGTTAACTTCGACCA | 225 |
| Qy | 225 | ATGCTCTCCCGAGCAGGTTCGATTATATACAGTCAATGGAAATGATAAAGCAAAATAGTTA | 284 |
| Db | 226 | GTGGACTCAAGAACAGATTTCAGTGCAATGCAAGAGATGGGAATGGAAAGGCNAACCGACT | 285 |
| Qy | 285 | CTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTGGAAATTGA | 329 |
| Db | 286 | TTATGAAGCCTATCTTCTCGAGACCTTTGGCGGACCTCAGATGGA | 330 |

```

RESULT 9
US-09-855-323-4
; Sequence 4, Application US/09855323
; Patent No. 6602667
; GENERAL INFORMATION:
; APPLICANT: Walker, Michael G.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: Inflammation-Associated Polynucleotides
; FILE REFERENCE: PB-0006-1 CIP
; CURRENT APPLICATION NUMBER: US/09/855,323
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 4
; .LENGTH: 1261
; .TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; .OTHER INFORMATION: 1221361CBI
US-09-855-323-4

Query Match          5.4%; Score 78.2; DB 4; Length 1261;
Best Local Similarity 60.7%; Pred. No. 2.6e-15;
Matches 128; Conservative 0; Mismatches 83; Indels 0; Gaps 0

Qy      119  GATCGGGCTAGTGTTAAATTIAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGA 178
Db      51  GATGGGGCTCTTGGAAACATTTGGTGTGTTTCATCTGCATTCGATGTGCTGGAATCCACAGGA 110

Qy      179  GTCTCGGGGTACACATATCGAAGGTTTCGATCTGCGCACTCTGGACACATGCGCTCCCGGAGC 238
Db      111  ATCTCGGGGTGACATATCCAGGTAAGTCAGTTAACTTCGACCACTGGACTCAAGAAC 170

Qy      239  AGGTTGCATTTATACAGTCAATGGGAATGATAAAGCAAAATAGTTTACTGGGAAGCAGAGC 298
Db      171  AGATTCAGTGCATCCTCAAGAGATGGGAATGGAAAGGCAAAACCGACTTTTATGAAGCCTATC 230

Qy      299  TACCCCAAACTATGATAGAGTTGGAAATTGA 329
Db      231  TTCTCTAGAGACCTTTGGCGGACCTCAGATAGA 261

```

```

RESULT 10
US-09-854-133-660
; Sequence 660, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedon
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Scrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0

```



```

; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 384
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 89..331
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 89..208
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq LNLGLVMCIETCSG/IH
US-09-513-999C-384

Query Match 4.7%; Score 68.4; DB 4; Length 332;
Best Local Similarity 57.5%; Pred. No. 2.le-12;
Matches 123; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 79 GAGAACAGAGATGTCTGACATGCAAAACAAAAGGTCCCAAGATGGGTAGTGTAAATTA 138
DB 119 GGAATCTCCACTGCGTGAGCTGTGAGACCCAGAAATCTTAAGTGGCCAGTTTGAATTG 178
QY 139 GGTATCTTTATCTGCAATGCAATGTTCTGGGATTTCACAGAGTCTCGGGGTACACATATCG 198
DB 179 GGAGTCTCATGTGTATTGAATGTTTCAGGAATCCACCGCAGTCTTGGCACCGCTTTCC 238
QY 199 AAGTTTCATCTGCCACTCTGGACACATGGTCCCGGAGCAGGTTCGATTTATACAGTCA 258
DB 239 CGTGTGCAATCTCTGGAGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 298
QY 259 ATGGGAATGATAAAGCAATAGTTACTGGGAAG 292
DB 299 ATTGCAATGACTAGCCAAACAGCATCCGGGAAG 332

RESULT 14
US-09-620-312D-35
; Sequence 35, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radjoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 35

; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 384
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 89..331
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 89..208
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq LNLGLVMCIETCSG/IH
US-09-513-999C-384

Query Match 4.7%; Score 68.4; DB 4; Length 332;
Best Local Similarity 57.5%; Pred. No. 2.le-12;
Matches 123; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 79 GAGAACAGAGATGTCTGACATGCAAAACAAAAGGTCCCAAGATGGGTAGTGTAAATTA 138
DB 119 GGAATCTCCACTGCGTGAGCTGTGAGACCCAGAAATCTTAAGTGGCCAGTTTGAATTG 178
QY 139 GGTATCTTTATCTGCAATGCAATGTTCTGGGATTTCACAGAGTCTCGGGGTACACATATCG 198
DB 179 GGAGTCTCATGTGTATTGAATGTTTCAGGAATCCACCGCAGTCTTGGCACCGCTTTCC 238
QY 199 AAGTTTCATCTGCCACTCTGGACACATGGTCCCGGAGCAGGTTCGATTTATACAGTCA 258
DB 239 CGTGTGCAATCTCTGGAGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 298
QY 259 ATGGGAATGATAAAGCAATAGTTACTGGGAAG 292
DB 299 ATTGCAATGACTAGCCAAACAGCATCCGGGAAG 332

RESULT 14
US-09-620-312D-35
; Sequence 35, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radjoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 35

; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (520) .. (1614)
; US-09-620-312D-35

Query Match 4.5%; Score 65.4; DB 4; Length 2119;
Best Local Similarity 56.9%; Pred. No. 8.7e-11;
Matches 120; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 82 AACAGAGATGTCTGACATGCAAAACAAAAGGTCCCAAGATGGGTAGTGTAAATTTAGGT 141
DB 898 AACAGCTTTTGTATCGAGTTCGATGACCAACCAATCCAGACTGGGCCAGCTGAACTGGGT 957
QY 142 ATCTTTATCTGCAATGTTCTGGGATTTCACAGAGTCTCGGGGTACACATATCGAAG 201
DB 958 GCCCTGATGTGATGTGAGTCTCAGGCATCCACCGACCTGGGGCTCACCTGTCTCCGG 1017
QY 202 GTTCGATCTGCCACTCTGGACACATGGTCCCGGAGCAGGTTCGATTTATACAGTCAATG 261
DB 1018 GTGCGCTCCCTTGACCTCGATGATGTCGCGCTGAGTGTGCTGTGCTGTGCTGTGCTGTG 1077
QY 262 GGAATGATAAAGCAATAGTTACTGGGAAG 292
DB 1078 GGCAATGCCCTCGCCAAACAGCGTCTGGGAGG 1108

RESULT 15
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9pt-Fls
US-08-232-463-14

Query Match      4.4%; Score 63.4; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 1e-09;
Matches 16; Conservative 228; Mismatches 149; Indels 0; Gaps 0;

Qy 264 AAATGATAAGCAATAGTTACTGGGAAGCAGAGCTACCCCAAACTATGATAGAGTTGG 323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1456 AAGAGATAGAGAAATTTGCTACRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1397

Qy 324 AATTGAGATTTTATACGTGCAAGTATCAAGAGAAGAGATGGTTCTTAGAGGGGAAAA 383
Db :: ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1396 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1337

Qy 384 GGCTAGATCACCTCCTAGAGTCAGCAGGACGCGGAAATCTGTGGAGAGAAGTGGGCC 443
Db :: ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1336 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1277

Qy 444 GGGATATGAGCATGGACATAGTAGTCTCTGTAAATTTGTTGAGGAGAGGAAAACTAT 503
Db ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1276 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1217

Qy 504 TCCAGCATAGACAAAGAAATAATGTTGCTGCAAGAGATAAATCTCCCGTCCTCC 563
Db :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1216 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1157

Qy 564 CCAAGGACCCAGTCAGGTTATAAGCCACAGCAGAAAAATGGAGTCTGCAGCTACTCCAGT 623
Db ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1156 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1097

Qy 624 AGAGAGGGAGAACACAGCAGTAAATGTTGCACC 656
Db ::::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1096 RFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFRFR 1064
```

Search completed: September 27, 2005, 06:22:54
Job time : 285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 27, 2005, 03:06:24 ; Search time 3611 Seconds
(without alignments)
2688.583 Million cell updates/sec

Title: US-10-630-518-1

Perfect score: 1452

Sequence: 1 atgaacgagaagcaacgt.....gaatttcacaaacattga 1452

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400732 seqs, 3343137571 residues

Total number of hits satisfying chosen parameters: 14801464

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/PT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/PTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------------|-------------------|
| 1 | 1452 | 100.0 | 1452 | US-10-630-518-1 | Sequence 1, Appli |
| 2 | 1450.4 | 99.9 | 1452 | US-10-630-518-3 | Sequence 3, Appli |
| 3 | 1450.4 | 99.9 | 1452 | US-10-630-518-5 | Sequence 5, Appli |
| 4 | 432.6 | 29.8 | 2640 | US-10-424-599-108715 | Sequence 108715, |
| 5 | 429.4 | 29.6 | 1879 | US-10-425-114-10983 | Sequence 10983, A |
| 6 | 348.6 | 24.0 | 2154 | US-10-767-701-14966 | Sequence 14966, A |
| 7 | 340 | 23.4 | 1951 | US-10-425-114-27747 | Sequence 27747, A |

| | | | | | | |
|----|-------|------|------|----|----------------------|--------------------|
| 8 | 340 | 23.4 | 2299 | 20 | US-10-425-115-176514 | Sequence 176514, |
| 9 | 338.4 | 23.3 | 2011 | 18 | US-10-425-114-4099 | Sequence 4099, Ap |
| 10 | 338.4 | 23.3 | 2022 | 18 | US-10-425-114-32446 | Sequence 32446, A |
| 11 | 336.6 | 23.2 | 1709 | 18 | US-10-425-114-22038 | Sequence 22038, A |
| 12 | 336.6 | 23.2 | 1802 | 18 | US-10-425-114-30303 | Sequence 30303, A |
| 13 | 336.6 | 23.2 | 1928 | 20 | US-10-425-115-176512 | Sequence 176512, A |
| 14 | 335 | 23.1 | 1915 | 18 | US-10-425-114-17368 | Sequence 17368, A |
| 15 | 333.4 | 23.0 | 2221 | 20 | US-10-425-115-176513 | Sequence 176513, |
| 16 | 331.8 | 22.9 | 1902 | 18 | US-10-425-114-2297 | Sequence 2297, Ap |
| 17 | 317.6 | 21.9 | 772 | 18 | US-10-424-599-84194 | Sequence 84194, A |
| 18 | 307 | 21.1 | 2180 | 19 | US-10-437-963-92697 | Sequence 92697, A |
| 19 | 298.2 | 20.5 | 596 | 19 | US-10-021-323-3493 | Sequence 3493, Ap |
| 20 | 255.4 | 17.6 | 5580 | 19 | US-10-630-518-8 | Sequence 8, Appli |
| 21 | 247.8 | 17.1 | 701 | 20 | US-10-425-115-24352 | Sequence 24352, A |
| 22 | 227.4 | 15.7 | 313 | 9 | US-09-294-093B-1317 | Sequence 1317, Ap |
| 23 | 227 | 15.6 | 779 | 20 | US-10-425-115-176517 | Sequence 176517, |
| 24 | 213.6 | 14.7 | 2291 | 18 | US-10-425-114-28016 | Sequence 28016, A |
| 25 | 213.6 | 14.7 | 2443 | 20 | US-10-425-115-176518 | Sequence 176518, |
| 26 | 206.6 | 14.2 | 1407 | 20 | US-10-739-930-3764 | Sequence 3764, Ap |
| 27 | 182.4 | 12.6 | 408 | 20 | US-10-856-499-1383 | Sequence 1383, Ap |
| 28 | 155.8 | 10.7 | 528 | 18 | US-10-424-599-105977 | Sequence 105977, |
| 29 | 155 | 10.7 | 732 | 20 | US-10-425-115-176137 | Sequence 176137, |
| 30 | 147.8 | 10.2 | 598 | 20 | US-10-425-115-176511 | Sequence 176511, |
| 31 | 127.2 | 8.8 | 290 | 9 | US-09-294-093B-2723 | Sequence 2723, Ap |
| 32 | 122.4 | 8.4 | 365 | 18 | US-10-424-599-136039 | Sequence 136039, |
| 33 | 116.8 | 8.0 | 388 | 20 | US-10-425-115-87733 | Sequence 87733, A |
| 34 | 107.2 | 7.4 | 587 | 9 | US-09-867-550-1487 | Sequence 1487, Ap |
| 35 | 107.2 | 7.4 | 2885 | 17 | US-10-120-988-3 | Sequence 1037, Ap |
| 36 | 107.2 | 7.4 | 3194 | 21 | US-10-887-553A-1037 | Sequence 170957, |
| 37 | 105.2 | 7.2 | 763 | 20 | US-10-425-115-170957 | Sequence 43, Appli |
| 38 | 98.4 | 6.8 | 1403 | 17 | US-10-426-776-43 | Sequence 3399, A |
| 39 | 95 | 6.5 | 1084 | 18 | US-10-425-114-31399 | Sequence 515, App |
| 40 | 94.2 | 6.5 | 1976 | 15 | US-10-037-270-515 | Sequence 515, App |
| 41 | 94.2 | 6.5 | 1976 | 17 | US-10-117-722-515 | Sequence 39539, A |
| 42 | 89.4 | 6.2 | 2263 | 20 | US-10-425-115-39539 | Sequence 52793, A |
| 43 | 85.2 | 5.9 | 649 | 20 | US-10-425-115-52793 | Sequence 1104, Ap |
| 44 | 82.6 | 5.7 | 1327 | 15 | US-10-037-270-1104 | Sequence 1104, Ap |
| 45 | 82.6 | 5.7 | 1327 | 17 | US-10-117-722-1104 | Sequence 1104, Ap |

ALIGNMENTS

RESULT 1
US-10-630-518-1
; Sequence 1, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFKY, Martin, F.
; TITLE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630.518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1452)
US-10-630-518-1

Query Match 100.0%; Score 1452; DB 19; Length 1452;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 1452; Conservative 0;

QY 1 ATGAACGAGAACCCACGCTCTTAAGGAGCTTAATGCGCGCATAGAAAGATTCTTGAA 60
DB 1 ATGAACGAGAACCCACGCTCTTAAGGAGCTTAATGCGCGCATAGAAAGATTCTTGAA 60
QY 61 GGGCTTCTTAACATCCAGACAGAGATGCTGACTGCAACAAAGAGTCCAGA 120
DB 61 GGGCTTCTTAACATCCAGACAGAGATGCTGACTGCAACAAAGAGTCCAGA 120
QY 121 TGGGCTAGTGTAAATTTAGTATCTTCTGATGCAATGTTCTGGGATTACAGGAGT 180
DB 121 TGGGCTAGTGTAAATTTAGTATCTTCTGATGCAATGTTCTGGGATTACAGGAGT 180
QY 181 CTCGGGGTACACATATCGAAGTTTCATCTGCACTCTGGAACATATGCTGCGAGCAG 240
DB 181 CTCGGGGTACACATATCGAAGTTTCATCTGCACTCTGGAACATATGCTGCGAGCAG 240
QY 241 GTTGCATTTATACAGTCAATGGGAATGATAAGCAATAGTTACTGGGAACGAGCTA 300
DB 241 GTTGCATTTATACAGTCAATGGGAATGATAAGCAATAGTTACTGGGAACGAGCTA 300
QY 301 CCCCCAACTATGATAGAGTTGGAATTTGAGAAATTTTATACGTGCAAGATATGAAGAG 360
DB 301 CCCCCAACTATGATAGAGTTGGAATTTGAGAAATTTTATACGTGCAAGATATGAAGAG 360
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTTAGAGTGGAGAACCGGGG 420
DB 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTTAGAGTGGAGAACCGGGG 420
QY 421 AAATCTGTGGAGAGATGGCGCGGATATGACATAGGACATAGTAGTCTGTAAAT 480
DB 421 AAATCTGTGGAGAGATGGCGCGGATATGACATAGGACATAGTAGTCTGTAAAT 480
QY 481 TTGTTTGAGGAGAGGAAACTATTCAGCATCTAGAAACAGAAATATGTTGCTGCAAG 540
DB 481 TTGTTTGAGGAGAGGAAACTATTCAGCATCTAGAAACAGAAATATGTTGCTGCAAG 540
QY 541 AGAATAAATCTTCCCGTGCTCCCCAAGGACCCAGTCAGGTTATAAGCCACAGAGAAA 600
DB 541 AGAATAAATCTTCCCGTGCTCCCCAAGGACCCAGTCAGGTTATAAGCCACAGAGAAA 600
QY 601 ATGGAGTCTGCAGTACTCCAGTAGAGGGGAGAACAGAGTAAATGTTCCACCAGCA 660
DB 601 ATGGAGTCTGCAGTACTCCAGTAGAGGGGAGAACAGAGTAAATGTTCCACCAGCA 660
QY 661 TCAGATCTCTCCAAAGGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
DB 661 TCAGATCTCTCCAAAGGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
QY 721 TCGACTACAAATACCTCAGAGGCAATCTCTGGGATACCTCTGCGGATGATTAATCTATGG 780
DB 721 TCGACTACAAATACCTCAGAGGCAATCTCTGGGATACCTCTGCGGATGATTAATCTATGG 780
QY 781 GCTGGCTTTTCAGTCTGCTGGAAGTGTCAACGGCAGAGAAAATGTCACGCCAGCCT 840
DB 781 GCTGGCTTTTCAGTCTGCTGGAAGTGTCAACGGCAGAGAAAATGTCACGCCAGCCT 840
QY 841 GCTGAGCAGATGTTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTTAAAGCACA 900
DB 841 GCTGAGCAGATGTTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTTAAAGCACA 900
QY 901 CCTAATTTAACTCAACAGCACAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
DB 901 CCTAATTTAACTCAACAGCACAACAAAGATGTGAAGGCGATATCATGAGCCTGTTT 960
QY 961 GAGAAGCAATATAGTATGCGCTTTTGCCATGTCACAGCAAGTTTCTATGCTCGCT 1020
DB 961 GAGAAGCAATATAGTATGCGCTTTTGCCATGTCACAGCAAGTTTCTATGCTCGCT 1020
QY 1021 CAGCAGAACCCCTTTTACATGGCTGAGGAGAAAGCTGCTGGAGGCACTCCAAACCGCGTG 1080

RESULT 2

US-10-630-518-3
; Sequence 3, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFKY, Martin, F.
; TITLE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630,518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1452)
US-10-630-518-3

Query Match 99.9%; Score 1450.4; DB 19; Length 1452;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCGCGCATAGAAAGATTCTTGAA 60
DB 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCGCGCATAGAAAGATTCTTGAA 60
QY 61 GGGCTTCTTAACATCCAGACAGAGATGCTGACTGCAACAAAGAGTCCAGA 120
DB 61 GGGCTTCTTAACATCCAGACAGAGATGCTGACTGCAACAAAGAGTCCAGA 120
QY 121 TGGGCTAGTGTAAATTTAGTATCTTATCTGCAATAGTCTTCTGGGATTCACAGGAGT 180

```
Db 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTACATGCAATGTTCTGGGATTTCACAGAGT 180
Qy 181 CTGGGGTACACATATCGAAGGTTGATCTGCCACTCTGACACATGGCTCCCGAGCAG 240
Db 181 CTGGGGTACACATATCGAAGGTTGATCTGCCACTCTGACACATGGCTCCCGAGCAG 240
Qy 241 GTTGCAATTTATACAGTCAATGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
Db 241 GTTGCAATTTATACAGTCAATGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
Qy 301 CCCCCAAACTATGATAGAGTTGGAATTTGAGAATTTTATACGTCCAAAGATGTAAGAAG 360
Db 301 CCCCCAAACTATGATAGAGTTGGAATTTGAGAATTTTATACGTCCAAAGATGTAAGAAG 360
Qy 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTTAGAGTCGAGCAGGACGGCGG 420
Db 361 AGATGGGTTTCTAGAGGGGAAAAGGCTAGATCACTCTTAGAGTCGAGCAGGACGGCGG 420
Qy 421 AAATCTGTGAGAGAGTGGGCGGGATATGAGCATGGACATAGTAGTCTCTGTAAAT 480
Db 421 AAATCTGTGAGAGAGTGGGCGGGATATGAGCATGGACATAGTAGTCTCTGTAAAT 480
Qy 481 TTGTTTTCAGGAGAGGAAAATTTTCCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 540
Db 481 TTGTTTTCAGGAGAGGAAAATTTTCCAGCATCTAGAACAAAGAAATATGTTGCTGCAACG 540
Qy 541 AGAATAAATCTTCGCTGCTCCCAAGGACCCAGTCAGGTTTAAAGCCACAGAGAAA 600
Db 541 AGAATAAATCTTCGCTGCTCCCAAGGACCCAGTCAGGTTTAAAGCCACAGAGAAA 600
Qy 601 ATGAGTCTCAGCTACTCCAGTAGAGAGGAGAGAAACAGCAGTAAATGTTGCACAGCA 660
Db 601 ATGAGTCTCAGCTACTCCAGTAGAGAGGAGAGAAACAGCAGTAAATGTTGCACAGCA 660
Qy 661 TCAGATCTCTCAAAGGTGGATTTTGCTACTGATCTGTTTAAACATGCTCAATGGAAT 720
Db 661 TCAGATCTCTCAAAGGTGGATTTTGCTACTGATCTGTTTAAACATGCTCAATGGAAT 720
Qy 721 TCAGTACAAATACCTCAGAGGCACTCTCCGCGATCTCTCCGATGATCTATCTATG 780
Db 721 TCAGTACAAATACCTCAGAGGCACTCTCCGCGATCTCTCCGATGATCTATCTATG 780
Qy 781 GCTGGCTTTTCAGTCTGCTGAGTGTCAACCGCAGAGAAATGTCACAGCAAGCCT 840
Db 781 GCTGGCTTTTCAGTCTGCTGAGTGTCAACCGCAGAGAAATGTCACAGCAAGCCT 840
Qy 841 GCTGAGAGCAGTTCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGCAC 900
Db 841 GCTGAGAGCAGTTCTCTCCAGCTTCACTCTTCTGACTTTGAGGATTTGTTTAAAGCAC 900
Qy 901 CCTAATTTTAACTAAAGCAGCAACCAAGATGTAAGGCGCATATCATGAGCCTGTTT 960
Db 901 CCTAATTTTAACTAAAGCAGCAACCAAGATGTAAGGCGCATATCATGAGCCTGTTT 960
Qy 961 GAGAGACGAATATAGTATCGCTTTTCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
Db 961 GAGAGACGAATATAGTATCGCTTTTCCATGATCAGCAACAGGTTGCTATGCTCGCT 1020
Qy 1021 CAGCAGCAAGCCCTTTTACATGGCTGAGCGGAAAGCTGCTGGAGGCACTCCAAACGGCGTG 1080
Db 1021 CAGCAGCAAGCCCTTTTACATGGCTGAGCGGAAAGCTGCTGGAGGCACTCCAAACGGCGTG 1080
Qy 1081 AATCAACAGCTATGCTAATGCTTTTAAAGCTAGCTTCTGCAATTTGGTCAAAACCCCGGC 1140
Db 1081 AATCAACAGCTATGCTAATGCTTTTAAAGCTAGCTTCTGCAATTTGGTCAAAACCCCGGC 1140
Qy 1141 GGCTACAGATCCCGGAATGCTAATGCTTTTAAAGCTAGCTTCTGCAATTTGGTCAAAACCTT 1200
Db 1141 GGCTACAGATCCCGGAATGCTAATGCTTTTAAAGCTAGCTTCTGCAATTTGGTCAAAACCTT 1200
Qy 1201 ATGCAAAACATGATGAAACGCAACATGCAACGAGAGCCCGCAACCGCAGAGAAC 1260
```

```
Db 1201 ATGCAAAACATGATATGAAACGCAAAACATGAAACGAGACCCGCAACCAACCGCAAGAAC 1260
Qy 1261 ACTCTACATATCCCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Db 1261 ACTCTACATATCCCATCCAGTTTCTACAAATGGGTCAAGCTAATCAAGTGAACGGT 1320
Qy 1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCAACCAAGCACC 1380
Db 1321 ATGACCCCAAACTCAACCGGTAAACCTCAGTCATCATCCGCAACCAACCAAGCACC 1380
Qy 1381 ACACATCTTCAATCAGGCAAGACTTTGATTTCTCTCTGATGATGGAATGTTTC 1440
Db 1381 ACACATCTTCAATCAGGCAAGACTTTGATTTCTCTCTGATGATGGAATGTTTC 1440
Qy 1441 ACAAACATTTGA 1452
Db 1441 ACAAACATTTGA 1452

RESULT 3
US-10-630-518-5
; Sequence 5, Application US/10630518
; Publication No. US20040143872A1
; GENERAL INFORMATION:
; APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: LILJEGREN, Sarah, J.
; APPLICANT: ECKER, Joseph, R.
; APPLICANT: YANOFKY, Martin, F.
; TITLE OF INVENTION: GENETIC CONTROL OF ORGAN ABSCISSION
; FILE REFERENCE: SALKINS.035C1
; CURRENT APPLICATION NUMBER: US/10/630,518
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: US 60/264,974
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCTUS02/01938
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(594)
US-10-630-518-5

Query Match 99.9%; Score 1450.4; DB 19; Length 1452;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGACGAGAAAGCCAACTCTCTTAAGGAGCTTAATGCCGCATAGAAAGATTCTTGAA 60
Db 1 ATGACGAGAAAGCCAACTCTCTTAAGGAGCTTAATGCCGCATAGAAAGATTCTTGAA 60
Qy 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGCTGACTGCAGAAACAAAAGTCCAAGA 120
Db 61 GGGCTTCTTAAACATCCAGAGAACAGAGAAATGCTGACTGCAGAAACAAAAGTCCAAGA 120
Qy 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTTCAAGAGT 180
Db 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTTCAAGAGT 180
Qy 181 CTCGGGGTACACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
Db 181 CTCGGGGTACACATATCGAAGGTTGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
Qy 241 GTTGCAATTTATACAGTCAATGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
Db 241 GTTGCAATTTATACAGTCAATGGAATGATAAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
Qy 301 CCCCCAAACTATGATAGAGTTGGAATTTGAGAATTTTATACGTCCAAAGATGTAAGAAG 360
```

Db 301 |||||CCCCAAACTATGATGAGTTGGAATTTAGAAATTTTATACGTCGCAAAAGTATGAAGAGAG 360
Qy 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTTAGAGTTCGAGCAGGAAACGGCGG 420
Db 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTTAGAGTTCGAGCAGGAAACGGCGG 420
Qy 421 AAATCTCTGAGAGAACTCGGCGGGATATGAGCATGGACATAGTAGTAGTCTCTGTAAAT 480
Db 421 AAATCTCTGAGAGAACTCGGCGGGATATGAGCATGGACATAGTAGTAGTCTCTGTAAAT 480
Qy 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGAACAGAAATATATGTTGCTGCAACG 540
Db 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGAACAGAAATATATGTTGCTGCAACG 540
Qy 541 AGAATAAATCTCCCGTGCTCCCAAGGAGCCAGTCAGGTTATAAGCCACAGCAGAAA 600
Db 541 AGAATAAATCTCCCGTGCTCCCAAGGAGCCAGTCAGGTTATAAGCCATAGCAGAAA 600
Qy 601 ATGGAGTCTGAGTACTCCAGTAGAGAGGGAGAAAACAGCAGTAAATGTTGCACGAGCA 660
Db 601 ATGGAGTCTGAGTACTCCAGTAGAGAGGGAGAAAACAGCAGTAAATGTTGCACGAGCA 660
Qy 661 TCAGATCTCTCAAAGTGGATTTTGCTGATCTGTTTAAATGCTATCAATGGATGAT 720
Db 661 TCAGATCTCTCAAAGTGGATTTTGCTGATCTGTTTAAATGCTATCAATGGATGAT 720
Qy 721 TCAGTACAATACTCCAGAGGCACTCTGGGGATCTCTGGCGATGATTAATCACTATGG 780
Db 721 TCAGTACAATACTCCAGAGGCACTCTGGGGATCTCTGGCGATGATTAATCACTATGG 780
Qy 781 GCTGGCTTTTCAGTCTGCTGAAAGTGGTCAAAACGGCAGAGAAAATGTCACAGCAGACCT 840
Db 781 GCTGGCTTTTCAGTCTGCTGAAAGTGGTCAAAACGGCAGAGAAAATGTCACAGCAGACCT 840
Qy 841 GCTGAGAGCAGTCTCTCCAGCTTCATCTCTGACTTTGAGGATTTGTTTAAAGCACA 900
Db 841 GCTGAGAGCAGTCTCTCCAGCTTCATCTCTGACTTTGAGGATTTGTTTAAAGCACA 900
Qy 901 CCTAATTTAACTCAACAGCACCACAAAGATGTAAGGGCGATATCATGAGCCTGTTT 960
Db 901 CCTAATTTAACTCAACAGCACCACAAAGATGTAAGGGCGATATCATGAGCCTGTTT 960
Qy 961 GAGAAGCAGTAATAGTATCGCTTTTGCCATGCATCAGCAACAGGTTGCTATGCTCGCT 1020
Db 961 GAGAAGCAGTAATAGTATCGCTTTTGCCATGCATCAGCAACAGGTTGCTATGCTCGCT 1020
Qy 1021 CAGCAGAACGCTTTTACATGGCTGACGGAAGGCTGCTGGAGGCACTCCAAACGGCGTG 1080
Db 1021 CAGCAGAACGCTTTTACATGGCTGACGGAAGGCTGCTGGAGGCACTCCAAACGGCGTG 1080
Qy 1081 AATCAACAGCTATGCTAATGCTCTTAACTGCTTAACTGCTTCTGCAAAATGGTCAAAACCCCGG 1140
Db 1081 AATCAACAGCTATGCTAATGCTCTTAACTGCTTAACTGCTTCTGCAAAATGGTCAAAACCCCGG 1140
Qy 1141 GGCTACAGATCCCGGAATGACTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAACTT 1200
Db 1141 GGCTACAGATCCCGGAATGACTAACCCCGTAGGTGGTCAAGCTGATCTCCAGAACTT 1200
Qy 1201 ATGCAAAAATGATATGAACGCAAAATGAACACGAGACCCGCAACACGCAAGAGAAC 1260
Db 1201 ATGCAAAAATGATATGAACGCAAAATGAACACGAGACCCGCAACACGCAAGAGAAC 1260
Qy 1261 ACTCTACAAATCCCATCATCTTCTACAAATGGTCAAGCTTCAAGTGAACGGT 1320
Db 1261 ACTCTACAAATCCCATCATCTTCTACAAATGGTCAAGCTTCAAGTGAACGGT 1320
Qy 1321 ATGACCCCAACTCAACCGGTAACCTCAGTCAATCAGCAGACCCCAACCAACAGCAC 1380
Db 1321 ATGACCCCAACTCAACCGGTAACCTCAGTCAATCAGCAGACCCCAACCAACAGCAC 1380
Qy 1381 ACACCAATTTCAAACTCAGGCAAGACTTTGATTTCTCTCTTGTGATGGAATGTTTC 1440
Db 1381 ACACCAATTTCAAACTCAGGCAAGACTTTGATTTCTCTCTTGTGATGGAATGTTTC 1440

Db 1381 ACACCAATTTCAAACTCAGGCAAGACTTTGATTTCTCTCTTGTGATGGAATGTTTC 1440
Qy 1441 ACACCAATTTGA 1452
Db 1441 ACACCAATTTGA 1452
RESULT 4
US-10-424-599-108715
; Sequence 108715, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108715
; LENGTH: 2640
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69183C.1
US-10-424-599-108715
Query Match 29.8%; Score 432.6; DB 18; Length 2640;
Best Local Similarity 61.5%; Pred. No. 3.2e-122;
Matches 914; Conservative 0; Mismatches 499; Indels 72; Gaps 11;
Qy 1 ATGACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAA 60
Db 215 ATGAAACAGCAAGGCCAAAGCTTTCCAAAGAGCTCAAGCCCAAGCAAGAGATACCTGAA 274
Qy 61 GGGCTCTTTAAACATCCAGAGAAACAGAGAAATGCTGACTGCAAAACAAAGAGTCCAAGA 120
Db 275 GGACTTCTTAAATACAGAGAAATAGGGAATGCTGACTGCAAGCTAAAGGTCCAAGA 334
Qy 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAT 180
Db 335 TGGGCTAGTGTAAATTTAGGTATCTTTATGCTGCAATGTTCTGGGATTCACAGGAT 394
Qy 181 CTCGGGTGACACATCAAGGTTCCGATCTGCACTCTGGACACATGGCTCCCGAGCAG 240
Db 395 TTGGGGGTACATATATCAAGGTTTCGTTCTGCAACCCCTGGACACTTGGCTTCCAGAGCA 454
Qy 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTGGGAAGCAGAGCTA 300
Db 455 GTTGCAATTTATCAATCAATGGGAAACGAGAAAGCAAAATGTTTCTGGGAAGCAATTA 514
Qy 301 CCCCCAAACTATGATAGAGTTGGAATTTGAGAAATTTTACGTGCAAGTATGAAGAGAG 360
Db 515 CTTCCAAATTTATGATAGAGTTGGAATTTGAGAAATTTCAATCTGCAAGTATGATGAAAG 574
Qy 361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTTCGAGCAGGAAACGGCGG 420
Db 575 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTTCGAGCAGGAAACGGCGG 420
Qy 421 AAATCTCTGAGAGAACTGGGCGGGATATGAGCATGGACATAGTAGTAGTCTCTGTAAAT 480
Db 632 AGTCTCTCACATTGGCAGAGGCTCTGGAGAGAGTGGTTATGCGCTGTTTCTGAAAT 691
Qy 481 TTGTTTGGAGAGAGAAAATCTATTCAGCATCTAGAACAGAAATATATGTTGCTGCAACG 540
Db 692 AAATTTGAGGAAAGAGAAAATCCAACTCACTGCAATTTCTACCGCAAGAAATTAAT 751
Qy 541 AGAATAAATCTTCCGTCGCTCCCCAAGGACCCAGTCAAGGTTTATAAGCCACAGCAGC 595
Db 752 GTTCTCTCTCTCCACAGCATCTGAGCAGGTAATCTCTTATTAACCAACCTCAGCATGTG 811

596 -AGAAATGGAGTCTGCAGTACTCAGTAGAGGGAGAAC-----AAGCAGTA 645
812 GAGAAAGTGGAAACCAAGTAGCAACCAACCAAGCTCCCAACCACTAGCTGAAACATCA 871
646 AATGTTGCACCAAGCATCAGAT-----CCTCAAAGGTGGATTTTGTACTACTGATCTG 696
872 AAGCAAGCTACAGACAGATTCAAAATATCCCTCCAAAGTTGACTATGCTACAGACCTT 931
697 TTTAATCATGCTATCAATGATGATTCGACTACAAATACCTCAGAGGAACTCCTGGCGAT 756
932 TTTCAACATGTTGCTATGATGATGCCCTAATGAAATGGCTCTGAGGCGAGCT-----GGT 985
757 ACTCTGCGGATGATTAATCATGGGCTGCTTTTCAAGTCTGCTGGAAGTGGTCAACCGCA 816
986 ACAAATCTAGTATGATTAATCATGGGCGAGTTTTCAGTCTGCTGCGAGGTGCAACAGCT 1045
817 GAGAAATTTGTCAACAGCAAGCTGCTGAGAGCAGTTTCTCTCCAGCTTCAATCTCTGA- 875
1046 GAGAGACTAGTCCCTCGAAGCAGCTGATGATCTCCAGGTTCTGCATCTGGAATTTGAG 1105
876 -----CTTTGAGGATTTTGAAGACACACCTAATTTAACTCAACAGCAACCAAAA 930
1106 GATCTTTTCAAGATTTTACATCTGTGACACCAAGTTTGACTCCAGAAAAACCAAGAA 1165
931 GATGTGAAGGCGATATCATGAGCTGTTTGAAGAGCAGCAATATAGTATCGGCTTTTGCC 990
1166 GATGTGAAGAAATGATATCATGAGCTCTTTTGAAGAGGCGCAATATGCTGCTCCANTTTCT 1225
991 ATGCATCAGCAACAGGTTGCTATGCTGCTCAGCAGCAAGCCCTTTTACATGG---CTGCA 1047
1226 ATGCATCAGCAGGCTTCCATGCTGCTAGCAGCAGCAAGCTCTCTTCTAATGCTTCTGCA 1285
1048 GCGAAAGCTGCTGGAGGCACTCAACCGGCGTGAATCAACAGCTATTGCTAATGCTCTT 1107
1286 GCTAAATCTTCTGGTGGGATCCTA-----GCTATCTGCTAGCATACACAACTTAGA 1339
1108 AAGTAGCTTCTGCAAAATTTGTTCAACCCCGGCGCTACAGATCCCGGATGACTTAAC 1167
1340 CCCAATGTTCTTATCCAAAGTTGGCCAGCTACTGGCTATTCAATCTCTGGAGT---AATG 1396
1168 CCGTAGTGGTCAAGCTGATCTCCAGAAATTTATGCAAAACATGAATATGAACGCAAC 1227
1397 CCATGGTGGTTCAGGAGAGTTGCAAAACCTTGTGCGAGGTTCAGACT-----1444
1228 ATGAACACGAGACCCGCAACCGCAAGAGAACACTCTTAATATCCATATCCTCAGTTTC 1287
1445 AGAACATGACCCGACACATTTTGCAGGAGTTCTGTTCAATATCCACCATCCAGTTTC 1504
1288 TACACAAATGGTCAAGCTAATCAAGTGAAGGTATGACCCCAACTCAACCGGTAAACCT 1347
1505 TATGGTATGGGCAAGTTCTCTCTCCAGTTAAGGTTATGACATGGGATGGATGAGTAACT 1564
1348 CAGTCATATCCGCAACCCCAACCAAGCAGCAGCAGCAGTCTTCAACATCAGGCAAGAC 1407
1565 CAGTCAGGAGCTCCTGAGTGTCTATCA-----CTTCAAGTCTGGAGGAT 1615
1408 TTTGATTTCTTCTTCTTGTGATGGAATGGAATGTTTCAACAAACATTTGA 1452
1616 TATGATTTTCTCTCTTAAACAGGGAATGTTTATGAACAGTGA 1660

RESULT 5

US-10-425-114-10983
; Sequence 10983, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10983
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700994349_FU1
US-10-425-114-10983

Query Match 29.6%; Score 429.4; DB 18; Length 1879;
Best Local Similarity 61.4%; Pred. No. 2.5e-121;
Matches 912; Conservative 0; Mismatches 501; Indels 72; Gaps 11;
QY 1 ATGAACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAAATGCCCCGCCATAGAAAGATTTTGA 60
DB 198 ATGAACGAGAAAGCCAAAGCTTTCAAAGAGCTCAACGCCAAAGCAACAGAGATATCTGAA 257
QY 61 GGGCTTCTTAAACATCCAGAAACAGAGAAATGCTGCTGACTGCAAAACAAAGAGTCCAAGA 120
DB 258 GGACTTCTTAAATTACCAGAGATAGGGAATGTGCTGACTGCAAAAGCTTAAGGTCCAAGA 317
QY 121 TGGCTAGTGTAAATTTAGGTATCTTTATCTGCAATGCAATGTTCTGGGATTCACAGGAGT 180
DB 318 TGGCTAGTGTAAATCTTGGCATCTTTATATGATGATGCTGTTTCAAGGAATACATCGAAGT 377
QY 181 CTCGGGTGACATATCGAAGGTTGATCTCCACTCTGGAACATCTGGACATAGTCTCCCGAGCAG 240
DB 378 TTGGGGGTACATATCAAGAGTTCGTTCTGCAACCCCTGGACACTTGGCTTCCAGAGCAA 437
QY 241 GTTGCAATTTACAGTCAATCGGAATGATAAGCAATAGTTACTTGGGAGCAGAGCTA 300
DB 438 GTTGCAATTTCAATCAATCGGAACCGAGAAACAAATGTTTCTGGAAAGCAATTA 497
QY 301 CCCCCAAACTATGATAGAGTTGGAAATTCAGAAATTTATACGTCGCAAGATATGAAGAGAG 360
DB 498 CTTCCAAATTTATGATAGAGTTGGAAATTCAGAAATTTTATCTGTCGCAAGATATGATGAAG 557
QY 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCTGAGCAGAGAAACGGCGG 420
DB 558 AGATGGGTTTCCAGAGAGATGGCAATTTCAAAACACCTTCTGGATT---TCGGGAAGAGAAA 614
QY 421 AAATCTGTGGAGAGAGTGGGCGGGATATGAGCATGACATAGTAGTAGTCTCTGTAAT 480
DB 615 AGTCTTCCACATTTGGCAGAGGCTGTGGAGAGAGTGGTTATGCCGCTGTTTCTGAAAT 674
QY 481 TTGTTTGAAGGAGAGAAATATTTCCAGCATCTAGAACAAAGAAATTAATGTTGCTGCAACG 540
DB 675 AAATTTGAGGAAGGAGAGAAATCCAACTCACTGCAATTCCTCTGCCCAAGAAATTAAT 734
QY 541 AGAATAAATCTTCCCGTCCCTCCCAAGGACCCAGTCCAGGTTTATAAGCCACAGC----- 595
DB 735 GTTCTGCTCTCTCCAGAGCACCTGAGCAGGTAATCTCTTATTTACCAAACTCTCAGCATGTG 794
QY 596 -AGAAATAGGAGTCTGAGAGTACTCCAGTAGAGAGGAGAAAC-----AAGCAGTA 645
DB 795 GAGAAAGTGGAAATCAGTAGCACCACCAACCAAGCTCCCAACCACTAGCTGAAACATCA 854
QY 646 AATGTTGCACCAAGCATCAG-----ATCCTCCAAAGGTGGATTTTGTCTACTGATCTG 696
DB 855 AAGCAGGCTACAGACACAGTTAAATTAACCTCTCTTAAAGTTGACTATGCTCCACAGACCTT 914
QY 697 TTTAATCATGCTATCAATGATGATTCGACTACAAATACCTCAGAGGCAACTCCTGGCGAT 756
DB 915 TTTCAACATGTTGCTCTATGATGGGCCCAATGAAATGGCTCTGAGGCGAGCT-----GGT 968
QY 757 ACTCTGCGGATGATTAATCATGGGCTGGCTTTTCACTGCTGCTGGAAGTGGTCAACCGCA 816
DB 969 ACAAATCTAGTATGATTAATCACTGGGCGAGTTTCCAGTCTGCTGCAAGGTGTCACAGCT 1028

817 GAGAAATGTCACAGCGAAGCTGCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGAC 876
1029 GAGAGACTAGTCCCTGAAAGCAGCTGATAGTACTCCAGGTTCTGCATCTGGAAATGAG 1088
877 -----TTTGAGGANTTTTAAAGGACACACCTAAATTTAAACAACTCAACAGCAACAAA 930
1089 GATCTTTTCAAAAGATTTACATCTGTGACACCAAGTTTGACTCCAGAAAACACAGAAA 1148
931 GATGTGAAGGGGATATCATGAGCCTGTTTGAGAGAGCAAGCAATATAGTATCGCCTTTGCC 990
1149 GATGTGAAGAAATGATATCATGAGCCTTTTGAGAGGGCAATATGTTCTCCATTTCT 1208
991 ATGCATCAGCAACAGGTTGCTATGCTGCTCAGCAGCAAGCCCTTTTACATGG---CTGCA 1047
1209 ATGCATCAGCAGCTTGCCATGCTAGCAGCAGCAAGCTCTCTTCTAATGCTTCTGCA 1268
1048 GGAAGAGCTGCTGGAGGCACTCAACAGCGGCTGAATCAACAGCTATTGCTAATGCTCTT 1107
1269 GCTAAATCTTCTGGTGGGATCCTA-----GGTATCTCTGTAGCATACAACAACCTAGA 1322
1108 AACGTAGCTTCTGCAAAATTTGGTCAAAACCGCGGCTACAGATCCCGGAATGACTAAC 1167
1323 CCCAATGTTCTTATCCAAAGTTGGCCAGCTACTGGCTATTCAATTCCTGGAGT---AATG 1379
1168 CCCGTAGGTGGTCAAGCTGATCTCCAGAAAATTTATGCAAAAACATGAATATGAACGAAAAC 1227
1380 CCCATGGTGGTTCAGGAGAGTTGCAAAACCTTTGTCAGGGTCAAGT----- 1427
1228 ATGAACACAGAGCCCGCAACACCGGAGAGAACACTCTCAATACCCATCATCAGTTTC 1287
1428 AGGAACATGACCCAGCAGCAATTTGAGGGAGTTCTGTTCAATATCAACATCCAGTTTC 1487
1288 TACACAAATGGGTCAAGCTAATCAAGTGAACGGTATGACCCCAAACTCAACCGGTAACT 1347
1488 TATGGTATGGGCAAGTTCTCTCTCAGTTAAGGTATGACAAATGAGGAGTGAATTAACCT 1547
1348 CAGTCATCATCCGCAACCCCAACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1407
1548 CAGTCAGGAGCTCTCCTCAGTGTCACTACCA-----CTTCAAGTCTCGGAAGGAT 1598
1408 TTGATTTCTCTCTCTGATGAGTGAATGTTTCACAAAACATGTA 1452
1599 TATGATTTTCTCTCTTAAACAGGGAATGTTTATGAACAGTGA 1643

RESULT 6

US-10-767-701-14966
; Sequence 14966, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767, 701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 14966
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS849_1
US-10-767-701-14966

Query Match 24.0%; Score 348.6; DB 19; Length 2154;
Best Local Similarity 60.3%; Pred. No. 2.8e-96;
Matches 645; Conservative 0; Mismatches 394; Indels 30; Gaps 3;
1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCCATAGAAAGATTTCTTGAA 60

RESULT 7

251 ATGAACGAGAAAGCGTCCGTTTCAAGGAGCTCAACGCCAAGCAACAAGATATTGGA 310
61 GGGCTTCTTAAACATCCAGAGAAAGAGAAAGTCTGACTCCAAACAAAGAGTCCAGA 120
311 AGTCTTCTTCCGCTTCTGAGAAAGAGAAAGTCTGCGGAACTGCAAGTCCAGAGGTCTCGA 370
121 TGGGCTAGTGTAAATTTAGGTATCTTATCTGCAATGCAATGTTCTGGGATTCACAGGAGT 180
371 TGGGCAAGTGTAAATCTAGGGATCTTATATATGATGACATGTTCTGGAAATTCATAGAAGC 430
181 CTCGGGTAACATATCGAAGTTTCGATCTGCACTCTGGAACATGCTGCTCCCGAGCAG 240
431 CTGGGGGTGCAATATCTAAGGTAAAGATCTGCAACCCCTGGGATACATGCTCCCGAGGCA 490
241 GTTCATTTATACAGTCAATGGGAAATGATAAAGCAATAGTTACTTGGGAAAGCAGAGCTA 300
491 GTTGCAATTTATCAATCAATGGGAAACGAAAGCAATAGCTATTGGGAAGCAGAGCTG 550
301 CCCCAAACTATGATAGAGTTGGAATTTGAGAAATTTATACGTGCAAAAGTATGAAGAGAG 360
551 CCTCTAATATGATAGGTTGGGATAGAGAAATTTCACTCCGTGCAAAATATAGGAGACAAG 610
361 AGATGGGTTCTAGAGGGGAAAAGGCTAGATCACTCTAGAGTCGAGCAGAGAAACGGCG 420
611 AGATGGGTACCGAGGAAATGGAAACATCAGACCTACCTCCGCTGTTCCGAGATGAGAAGAGC 670
421 AAATCTGTGGAGAGAGTGGCCCGGATATGAGCATAGGACATAGTAGTAGTCTCTTAAT 480
671 CAAGAGTCTCTGGCTAGTGTCTAATAGGAGTGAACATGCTCATAGATCTTCAATTTGAGCA 730
481 TTGTTTGAGGAGAGAGAAAACCTATTTCAGCATCTAGAACAAAGAAATAATTTGCTGCAACG 540
731 AACCGTTTACAGCTCTTCGAGCAAGTTGCACTGTAGCTTCAAGGATACCTCTCTCAG 790
541 AGAATAAATCTTCCGTCCTCCCAAGGACCCAGTCAAGTTTAAAGCCACAGCAGAAA 600
791 GCATCACTCAGCCGCCCAAGAGTGAACACCACTTCTAAGGTGGTTTCACTCTCTCAG 850
601 ATGAGTCTGAGCTACTTCCAGTAGAGAGGAGAAACAAAGCAGTAAATTTGTCACACAGCA 660
851 CCACAGAAATCTCTGCAAAAGTTGAGGCAACACCCCTAAAGTTGAGAAGCCATCAGTT 910
661 TCAGATCTCCAAAGGTGGATTTTGTCTACTGATCTGTTTAACTGCTATCAATGATGAT 720
911 GCGCGCTCTCTAAAGTCGATTATGCTACCGATCTCTTAAACATGTTATCAATGGATGGA 970
721 TCGACTACAATATCTCAGAGGCAACTCTCTGGCGATACTCTCGCATGATAACTCATGG 780
971 ACAACAGAGAAAGAGTCAGAGTCACTCTTCAAAT-----GATGATAATGCCCTGG 1018
781 GCTGCTTTTCACTGCTGCTGGAAGTGGTCAACCGCAGAGAAAATTTGTCAAGCCCAAGCT 840
1019 GATGATTTCCAGTCTGCAACACCACTAGCTAGCTCAGAGAAAAGAGATTTCTGCCAACC 1078
841 GCTGAGAGCAGTCTCTCTCCAGCTTCACTTCTGACTTTGAGGATTTGTTTAAAGACACA 900
1079 GCAGAAAGCAAGACCC-----AGTCAACATCCGGAATAGAAGACCTATTTTAAAGACTCA 1132
901 CCTAATTTAACAACATCAACAGCACC-----AAAAGATGTCAAGGGCGATATC 948
1133 CCAGTCTGTCAATATCTCTCAGCTCCAGCTGTTTCCCAAGTAAATGCGAAGATGATATC 1192
949 ATGAGCCTGTTTGAGAAAGCAATATAGTATCGCTTTTGGCCATGCTATCAGCAACAGGTT 1008
1193 ATGAGTTTGTGAGAAAGTCCATATGTTATGTTATCGCTTCCGCTCCCATCAACAGCAGCTG 1252
1009 GCTATGCTGCTCAGCAGCAAGCCCTTTTACATGCTGCGAGGAAAGCTG 1057
1253 GCGTTTATGCTCAGCAGCAAGCTCTTCTAATGCTGCTCTTAAAGCTG 1301

US-10-425-114-27747
; Sequence 27747, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27747
; LENGTH: 1951
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4739-018-P5_FLI
US-10-425-114-27747
Query Match 23.4%; Score 340; DB 18; Length 1951;
Best Local Similarity 61.0%; Pred. No. 1.2e-93;
Matches 654; Conservative 0; Mismatches 385; Indels 33; Gaps 5;
QY 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGA 60
DB 259 ATGAACGAGAGGCGTCTCTCCAGGAGCTCAAGCCCAAGCACAGAAGATATTGGA 318
QY 61 GGGCTTTTAAATCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAGGTCAGA 120
DB 319 GGTCTTCTACGGCATCTCTGAGAAATAGAAATGCGAGACTGCAAGTCAAGGGTCTCGA 378
QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT 180
DB 379 TGGCAAGTGTGAATCTAGTATCTTTATGATGATGACATGTTCTGGCATTCATAGAGC 438
QY 181 CTCGGGGTACACATATCGAAGTTGATCTGCACTCTGGACACATGCTCCCGAGCAG 240
DB 439 CTGGGGGTGACATATCTAAGTAAAGTCTGCCACCTGGATACATGCTGCCAGCA 498
QY 241 GTTGCAATTATACAGTCAATGGGAATGATTAAGCAATAGTACTTGGGAACAGAGCTA 300
DB 499 GTTGCAATTATTAATCAATGGGAACGAAAAAGCAAAATAGCTATTGGGAAGCAGAGCTG 558
QY 301 CCCCACAACTATGATAGTGTGMAATTTGAGAAATTTATACGTGCAAAAGTATGAAGAGAG 360
DB 559 CTTCTTAATCTAGATAGGTTGGAATAGAGAAATTTTATCCGTGCAAAATATGAGGACAG 618
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTCCAGCAGGAACGGCGG 420
DB 619 AGATGGGTACCAAGGATGGAACATCAAAATCTTCGTCCAGTTCGAGATGAGAAGAGC 678
QY 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGGAATAGTAGT--AGTCTGTAA 478
DB 679 CAAGAGTCTCCGGCTAGTGTAAATAGGAGTGGACATGGTTCATAGATCTTTCAATTTGAGCAA 738
QY 479 ATTTGTTTTCAGGAGGAGAAACATATTCAGCATCTAGACAGAAATATGTTGCTGCAA 538
DB 739 AACCGTGTTCACCACTCTTCGAGCAAGTTGCAATGCGAGCTTCAAGGATATCTCT 798
QY 539 CGAGAATAAA-TCTTCCCTGCTCCCAAGGACCCAGTCAGGTTATAAAGCCACAGCAG 597
DB 799 CAGGCATACCTCAGCCCCCGAGGTAGAAACACAGTTCCAAAGGTGTTTACTCTCT 858
QY 598 AAAATGGAGTCTGACGCTACTCCAGTGAAGAGGAGAAACAGAGCAATAATGTTGACCA 657
DB 859 CAGTCAAGAAATCCCCCGCAAGGTTGATGCAACACACCGCTAAAGGTTGAGAAGCCATCG 918
QY 658 GCATCAGATCTCCAAAGGTGGATTTTGCTACTGATCTGTTTAAACATGCTATCAATGGAT 717

DB 919 GTTACACACACCTCTCTAAAGTTGATTATGCCACTGATCTCTTTAAACATGTTGTCAATGGAT 978
QY 718 GATTTCGACTACAAATACCTCAGAGGCAACTCTCTGGGATATCTCTGCCGATGATACTCA 777
DB 979 GGAACAAACAGAGAAAGAGTCAAGCTCA-----TCTTCAACAGCAGATTAATGGC 1026
QY 778 TGGGCTGGCTTTTCACTCTGCTGGAAAGTGGTCAAAAGCGCAGAGAAAAATTTGTCAACGCCAAG 837
DB 1027 TGGGATGGCTTCCAGTCTGCAACACCACTAGTACTTTCAGAGAAAAAGATTCTGCCAAA 1086
QY 838 CTTCTCAGAGCAGTCTCTCTCCAGCTTCTCTCTGACTTCTGAGTATGAGGATTTGTTTAGAC 897
DB 1087 CCAGCAGAAAGCAAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAAGAC 1140
QY 898 ACACCTAATTTAACTCAACAGCACCACCAAGATGGAAGGC-----GAT 945
DB 1141 TCACAGCTGTGCCATTATCTCAGCTCCAGCTGTTTCCCAGTAAACGCTAAGATGAT 1200
QY 946 ATCATGAGCCTGTTTGAAGAGCAATATAGTATCGCTTTTGGCATGATCAGCAACAG 1005
DB 1201 ATCATGAGTGTGTTTGAAGAGTCCCAATATGATATCACCTTTGCTGCCCATCAACAGCAG 1260
QY 1006 GTTCTATGCTGCTCAGCAGCAAGCCCTTTTACATGCTGCGCAGGAAAGCTG 1057
DB 1261 CTGGCGCTCATGCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTCAAAGCTG 1312
RESULT 8
US-10-425-115-176514
; Sequence 176514, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176514
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92573C.1
US-10-425-115-176514
Query Match 23.4%; Score 340; DB 20; Length 2299;
Best Local Similarity 61.0%; Pred. No. 1.4e-93;
Matches 654; Conservative 0; Mismatches 385; Indels 33; Gaps 5;
QY 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGA 60
DB 327 ATGAACGAGAAAGGCTCGTCTCCAGGAGCTCAAGCCCAAGCACAGAGATATTGGAA 386
QY 61 GGGCTTTTAAACATCCAGAGAACAGAGAAATGTGCTGACTGCAAAACAAAGGTCAGA 120
DB 387 GGTCTTCTACGGCATCTCTGAGAAATAGAAATGCGCAGACTGCAAGTCAAGGGTCTCTCGA 446
QY 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT 180
DB 447 TGGCAAGTGTGAATCTAGGATCTTTATATGATGATGATGTTCTTGGCATTCATGAAGC 506
QY 181 CTCGGGGTACACATATCGAAGTTCGATCTGCCATCTGGAACACATGCTGCCGAGCAG 240
DB 507 CTGGGGGTGACATATCTAAGGTAAAGTCTGCCACCTTGGATACATGCTGCCAGGCAA 566
QY 241 GTTGTATTTATACAGTCAATGGGAAATGATTAAGCAAAATAGTACTTGGGAAGCAGAGCTA 300

Db 567 GTTGCAATTATTCAATCAATGGGAACGAAAGCAAAATAGCTATTGGGAAGCAGAGCTG 626
Qy 301 CCCCCAACTATGATAGAGTTGGATTTGAGATTTTATACGTGCAAGATGATGAAGAGAG 360
Db 627 CTTCTTAATCATGATAGAGTTGGAATAGAGATTTTATCCGTGCAAAATATAGAGCAAG 686
Qy 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACCTCTCTAGAGTGCAGAGCAAGCGGG 420
Db 687 AGATGGGTACCAAGGATGGAACATCAAAATCTTCGTCCAGTGTTCGATGAGAGAGC 746
Qy 421 AAATCTGTGGAGAGAGTGGCGGGAATAGAGCATAGGACATAGTAGT--AGTCTGTAA 478
Db 747 CAAGAGTCTCCGGCTAGTGTAAATAGAGTGCACATGGTCTATAGATCTTCATTTGAGCAA 806
Qy 479 ATTTGTTTGGAGGAGGAAATATATCCAGCATCTAGACAGAAATATATGTTGCTGCAA 538
Db 807 AACCGTGTCTCCAGCTCTCCGAGCAAGTTGCACATGCGAGCTTCAAGGATATCTCT 866
Qy 539 CGAGAAATAA--TCTTCCCGTGTCTCCCAAGGACCCAGTCCAGTATATTAAGCCACAGCAG 597
Db 867 CAGGCATCACCTCAGCCCCCGAAGGTAGAAACACCAAGTTCCAAAGTGGTTTACCTCT 926
Qy 598 AAATGGAGTGTGAGTACTCAGTAGAGAGAGGAGAAACAGAGCATAGTAATGTTGCACCA 657
Db 927 CAGTCACAGAAATCCCCCGCCAAAGTTGATGCAACACCCGCTTAAAGTTGAGAAGCCATCG 986
Qy 658 GCATCAGATCTCCAAAGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGGAT 717
Db 987 GTTACACCACTCTCTAAAGTTGATATGCACTGATCTCTTTAAACATGTTGTCATGGAT 1046
Qy 718 GATTGCACTACAAATACCTCAGAGGCAACTCTGCGCATACTCTCTGCGGATGATTAATCTCA 777
Db 1047 GGAACAACAGAGAGAGTCAAGTCA-----TCTTCAACGAGCATATGGC 1094
Qy 778 TGGGTGGCTTTTCACTGTCTGGAAGTGTGTAACGGGAGAGAAATTTGTACAGCCCAAG 837
Db 1095 TGGGATGGCTTCCAGTCTGCAACACAGTACCTAGTTTCAGAGAAAAAGATTTCTGCCAAA 1154
Qy 838 CTGTGTGAGAGAGTTCTCTCCAGTTCATCTCTGATTTTGGAGATTTGTTTAAAGAGC 897
Db 1155 CCAGCAGAAAGCAAGACCCAG-----TCTATCTTGGATGGAAGACTTATTTAAAGAC 1208
Qy 898 ACACCTAATTTAACTCAACAGCACCAAGAGATGTGAAGGC-----GAT 945
Db 1209 TCACCACTGTGCCATTATCTCTCAGCTCCAGCTGTTTCCAGTAAACGCTAAGATGAT 1268
Qy 946 ATCATGAGCTTTTGAAGAAGCAATATAGTATGCGCTTTTGGCCATGCTATGAGCAAG 1005
Db 1269 ATCATGAGTTTGTGAGAAGTCCAATATGATGATCACCTTTGCTGCCCCATCAACAGCAG 1328
Qy 1006 GTTGCTATGCTCGCTCAGCAGACGCCCTTTATCATGGCTGCGAGCAAGAGCTG 1057
Db 1329 CTGGCGCTCATGTCTCAGCAGCAGCTCTTCTAATGGGTGCTCTCAAGAGCTG 1380

RESULT 9

US-10-425-114-4099
; Sequence 4099, Application US/10425114
; Publication NO. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nuclear Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4099

LENGTH: 2011
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
; OTHER INFORMATION: Clone ID: 700349043_FLI
US-10-425-114-4099
Query Match 23.3%; Score 338.4; DB 18; Length 2011;
Best Local Similarity 60.9%; Pred. No. 3.9e-93;
Matches 653; Conservative 0; Mismatches 386; Indels 33; Gaps 5;
Qy 1 ATGAACGAGAAAGCAACGCTCTCTAAGGAGCTTAATGCCGCCATAGAAAGATTCTTGAA 60
Db 170 ATGAACGAGAAAGCGTCCGTCTCAAGGAGCTCAAGCCCAAGCAACAAGATATTGAA 229
Qy 61 GGGCTTCTTAACATCCAGAGAACAGAGAAATGTCTGACTGCAAAACAAAAGGTCGAAGA 120
Db 230 GGTCTTCTACGGCATCTCGAATAGAGAAATGCCAGACTGCAAGTCAAGGGTCTCGA 289
Qy 121 TGGGCTAGTGTAAATTTAGGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGGAGT 180
Db 290 TGGCAAGTGTGAATCTAGGGATCTTTATATGATGATGACATGTTCTGGCATTCATAGAGC 349
Qy 181 CTCGGGTACACATATCGAAGGTTGATCTGCCACTCTGGAACATATGGCTCCCCGAGCAG 240
Db 350 CTGGGGGTGCACATATCTAAGGTAAAGATCTGCCACCTCGGATACATGGCTGCCAGAGCAA 409
Qy 241 GTTCATTTTATACAGTCAATGGGAATGATAAGCAATAGTCTTACTGGGAAGCAGAGCTA 300
Db 410 GTTCATTTTATCAATCAATGGGAACGAAAGCAATAGTCTTTGGGAAGCAGAGCTG 469
Qy 301 CCCCCAACTATGATAGAGTTGGAAATTTTATACGTGCAAAAGTATGAAGAGAG 360
Db 470 CTTCTTAATCATGATAGAGTTGGATAGAGATTTTATCCGTGCAAAATATGAGGACAG 529
Qy 361 AGATGGGTTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTGCAGAGCAAGCGGG 420
Db 530 AGATGGGTACCAAGGAATGGAACATCAAAATCTTCTGCTCAGTGTTCGAGATGAGAAGAGC 589
Qy 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGAGCATAGTAGT--AGTCTGTAA 478
Db 590 CAAGAGTCTCCGGCTAGTGTAAATAGAGTGGACATGTCATAGATCTTCAATTTGAGCAA 649
Qy 479 ATTTGTTTGGAGAGAGGAAACTATTCAGCATCTAGAAACAAGAAATAATGTTGTCGAA 538
Db 650 AACCGTCTTCGCCAGCTCTTCGAGCAAGTTGCAATGCAAGTTCAGGATATCTCT 709
Qy 539 CGAATAATAA--TCTTCCCGTGCCTCCCAAGGACCCAGTCAAGTTATTAAGCCACAGCAG 597
Db 710 CAGGCATCACCTCAGCCCCCGAAGGTAGAAACACCAAGTTCCAAAGGTGTTTACCTCT 769
Qy 598 AAATGGAGTCTGCAGCTACTCCAGTAGAGAGGAGAAACAAGCAGTAAATGTTGCACCA 657
Db 770 CAGTCACAGAAATCCCCCGCAAGTTGATGCAACACCGCTTAAAGTTGAGAGCCATCG 829
Qy 658 GCATCAGATCTCCAAAGTGGATTTTGTACTGATCTGTTTAACTGCTATCAATGAT 717
Db 830 GTTACACCACTCTCTAAAGTTGATTTATGCCACTGATCTCTTTAAACATGTTGTCATGAT 889
Qy 718 GATTGCACTACAAATATCTCAGAGGCAACTCTCTGGGATACTCTCTGCGATGATTAATCTCA 777
Db 890 GGAACAACAGAGAAAGAGTCAAGTTCA-----TCTTCAAGCAGCATTAATGGC 937
Qy 778 TGGGCTGCTTTTCAAGTCTGCTGGAAAGTGGTCAAAACGGCAGAGAAATTTGTCAAGCCCAAG 837
Db 938 TGGATGGCTTCCAGTCTGCAACACAGTACCTAGTTGAGAGAAAAAAGATTTCTGCCAAA 997
Qy 838 CTTCTGAGAGCAGTCTCTCTCCAGCTTCTATCTTCTGATTTGAGGATTTGTTTAAAGAC 897
Db 998 CCACGAGAAAGCAGAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAAGAC 1051
Qy 898 ACACCTAATTTAACTCAACAGGACCAAGCATGTGAAGGC-----GAT 945

Db 1052 TCACGAGTGTGCCATTATCTCCAGCTCCAGCTGTTTCCCAAGTAAACGTAAGAATGAT 1111
Qy 946 ATCATGAGCTGTTTCAGAGAGCAATATAGTATGCGCTTTTGCCATGCGTCAAGCAAG 1005
Db 1112 ATCATGAGTGTGTTGAGAAGTCCAAATATGGTATCACTTTCGCTGCCCATCAAGCAG 1171
Qy 1006 GTTGTCTATGCTCGCTCAGCAGCAAGCCCTTTACATGGCTGCGAGCAAGAGCTG 1057
Db 1172 CTGGCGCTCATGTCTCAGCAGCAAGCTCTTTCAATGGCTGCTCTCAAGCTG 1223

RESULT 10
US-10-425-114-32446
; Sequence 32446, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32446
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73346C08_FLI
US-10-425-114-32446

Query Match 23.3%; Score 338.4; DB 18; Length 2022;
Best Local Similarity 60.9%; Pred. No. 3.9e-93;
Matches 653; Conservative 0; Mismatches 386; Indels 33; Gaps 5;

Qy 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTCTTGAA 60
Db 242 ATGACGAAAGAGGCTCGCTCTCCAGGAGCTCAACGCCAAGCACAAGAAAGATATTGGAA 301
Qy 61 GGGCTTTTAAACATCCAGAGAACAGAGAAATGTCTGACTGTCAAAACAAAGGTCACAGA 120
Db 302 GGTCTTTACGGCATCCTCGAGAATAGAGAATGCGCAGACTGCAAGTCAAAAGGGTCTCTGA 361
Qy 121 TGGGCTAGTGTAAATTTAGTATCTTTATCTGATGCAATGTTCTGGGATTCACAGGAGT 180
Db 362 TGGGCAAGTGTGAATCTAGGTATCTTTATGATGATGACATGTTCTGGCATTCATAGAGC 421
Qy 181 CTCGGGTTACACATATCGAAGTTCGATCTGCACTCTGACACATGCTCCCGCAGCAG 240
Db 422 CTGGGGTGCACATATCTAAGTAAATCTGCCACCTGATACATGCTGCCAGCAGCAA 481
Qy 241 GTTGCAATTTATACAGTCAATGGAATATATAAGCAAAATAGTTACTTGGGAAGCAGACTA 300
Db 482 GTTGCAATTTATCAATCAATGGAACGAAAGCAATAGTATTTGGGAAGCAGAGCTG 541
Qy 301 CCCCACAACTATGATAGATTGGAAATGGAATTTATACGTGCAAGTATGAGAGAG 360
Db 542 CTTCTTAACATGATAGAGTTGGAATAGAGAATTTTATCCGTGCAAAATATGAGGACAG 601
Qy 361 AGATGGGTTCTTAGAGGGGAAAGGCTAGATCACTCTTAGAGTCAGCAGGAAGCGCGG 420
Db 602 AGATGGGTACCAAGGAATGGAACATCAAAATCTTCTGTCAGTGTTCGATGAGAGAGC 661
Qy 421 AAATCTGTGGAGAAAGTGGCGCGGATGAGCATGAGCATGAGTAGT--AGTCTGTAA 478
Db 662 CAAGAGTCTCCGCTAGTGTAAATAGAGTGGACATGGTCAATAGATCTTCAATTTAGCAA 721
Qy 479 ATTGTTTGGAGGAGGAAACATATTCAGCATCTAGAACAGAAATATGTTGTGCAA 538

Db 722 AACCGTGTCTCACCAGCTCTTCCGAGCAAAAGTTGCATGCGCTTCAAGGATATCTCT 781
Qy 539 CGAAGATAAA-TCTTCCGTCCTCCCAAGGAGCCAGTCAAGTTATAAAGCCAGCAG 597
Db 782 CAGGCATCACCTCAGTCCGCCCCGAAAGTAGAAACACAGTTTCAAAGGTGGTTTCTCTCT 841
Qy 598 AAATGGAGTCTGCGAGTCTACTCCAGTAGAGAGGAGAAACAAAGCAGTAAATGTTGCACCA 657
Db 842 CAGTCACAGAAATCCCCCGCCAAAGTTGATCAACACCCCTAAAGTTGAGAAGCCATCG 901
Qy 658 GCATCAGATCTCCAAAGGTGGATTTTGTACTGATGCTTGTAAATGCTATCAATGAT 717
Db 902 GTTACACCACCTCTCTAAAGTTGATATGCCACTGATCTCTTTAAACATGTTGTCAATGGAT 961
Qy 718 GATTGCACTACAAATACCTCAGAGGCACTCTCGCGGATACCTCTGCGGATGATAACTCA 777
Db 962 GGAACACAGAGAAAGAGTCAGCGTCA-----TCTTCAACAGCATTAATGGC 1009
Qy 778 TGGGCTGGCTTTTCAGTCTGTGGAAGTGGTCAACGGCAGAGAAATTTGTCAAGCCCAAG 837
Db 1010 TGGGATGGCTTCAGTCTGCAACACCACTAGTTCAGAGAAAAAGATTCTGCCAA 1069
Qy 838 CTGCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGACTTTGAGGATTTGTTTAAAGGAC 897
Db 1070 CCAGCAGAAAGCAAGACCCAG-----TCTACATCTGGAATGGAAGACTTATTTAAAGAC 1123
Qy 898 ACACCTAATTTAAACATCAACAGCACCACCAAGATGTGAAGGC-----GAT 945
Db 1124 TCACGAGCTGTGCCATTAATCTCAGCTCCAGCTGTTTCCCAAGTAACGCTAAGATGAT 1183
Qy 946 ATCATGAGCTGTTTGAGAGACGAATATAGTATGCGCTTTTGCCATGCGTCAAGCAAG 1005
Db 1184 ATCATGAGTGTGTTGAGAGTCCAAATATGATGATCACTTTCGCTGCCCATCAAGCAG 1243
Qy 1006 GTTGTATGCTGCTCAGCAGCAAGCCCTTTACATGCTGCGTGCAGCGAAGCTG 1057
Db 1244 CTGGCGCTCATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTCAAGCTG 1295

RESULT 11
US-10-425-114-22038
; Sequence 22038, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22038
; LENGTH: 1709
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-022-P10_FLI
US-10-425-114-22038

Query Match 23.2%; Score 336.6; DB 18; Length 1709;
Best Local Similarity 58.9%; Pred. No. 1.3e-92;
Matches 626; Conservative 0; Mismatches 419; Indels 18; Gaps 2;

Qy 1 ATGAACGAGAAAGCCAACTCTCTAAGGAGCTTAATGTCGCGCCATAGAAAGATTCTTGAA 60
Db 77 ATGAACGAGAAAGGCGTCCGCTCTCAGGAGCTCAACGCCAAGCAAGAAAGATATTGGAA 136


```
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3069-017-C3_FLI
US-10-425-114-17368

Query Match      23.1%; Score 335; DB 18; Length 1915;
Best Local Similarity 58.8%; Pred. No. 4.3e-92;
Matches 625; Conservative 0; Mismatches 420; Indels 18; Gaps 2;

QY 1 ATGACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTTCTTGAA 60
DB |||||
QY 277 ATGACGAGAAAGCGTCCGCTCTCAAGAGCTCAAGCCAGCAAGAGATATTGGAA 336
DB |||||
QY 61 GGGCTTTCTTAAACATCCAGAGAAACAGAGATGTGCTGACTGCAAAACAAAAGTCCCAAG 120
DB |||||
QY 337 GGTCTTTCTCGGCTTCTGAGATAGAGATGTGCAGACTGCAAGTCAAAGGTCTCTCGA 396
DB |||||
QY 121 TGGGCTAGTGTTAATTTAGTATCTTTATCTGATGCAATGTCTGGGATTTACAGAGT 180
DB |||||
QY 397 TGGGCAAGTGTTAATCTCGGATCTTTATATGTATGACATGTTCTTGGAAATTCATAGAAGC 456
DB |||||
QY 181 CTCGGGTTACACATATCGAAGTTCGATCTGCCACTCTGCACATGGCTCCCGAGCAG 240
DB |||||
QY 457 CTGGGGTGCACATATCTAAGGTAAAGTCTGTACCTCGATACATGCTGGCTGCCAGAGAA 516
DB |||||
QY 241 GTTGCAATTTATACAGTCAATGGGAAATGATAAGCAAAATAGTTACTTGGGAAGCAGAGCTA 300
DB |||||
QY 517 GTTGCAATTTATCCAAATCAATGGGAAACGAAAGCAAAATAGCCATTGGGAAGCAGAGCTG 576
DB |||||
QY 301 CCCCCAACTATGATAGAGTTGGAATTTATAGATTTTATAGTGCAAAAGTATGAAGAGAG 360
DB |||||
QY 577 CTTCTTAACTATGATAGAGTTGGGATAGAGAAATTTCTCCGTGCAAAATATGAGGACAG 636
DB |||||
QY 361 AGATGGTCTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGAGTGCAGCAGGAACGGGG 420
DB |||||
QY 637 AGATGGGTACCGAGGAATGGAACATTCGGACCTTCTCCGGTTCGAGATGAGAGAGC 696
DB |||||
QY 421 AAATCTGTGGAGAGAGTGGCGGGATATGAGCATGAGCATAGTAGTAGTCTCTGTAAT 480
DB |||||
QY 697 CAAGAGTCTCCAGTAGTGTCTAATAGAGTGGAGCATCATAGATCTTCATTTGAGCAAAAC 756
DB |||||
QY 481 TTGTTTGAGAGAGGAGAAACTATTCAGCATCTAGAACAGAGAAATATGTTGCTGCAAGC 540
DB |||||
QY 757 CGTGCTTCCAGCTCTCCGCAAGTTGAGCAATGTGCAATGTAGTCTCAAGGATACCTCTCAG 816
DB |||||
QY 541 AGAATAAATCTTCCGCTGCTCCCAAGGACCCAGTCAGGTTATAAAGCCACAGCAGAA 600
DB |||||
QY 817 GCATCACCTCAGCCCCCAAGGTAGAACCCACAGTTCTTAAGGTGTTCACTCTCAG 876
DB |||||
QY 601 ATGGAGTCTGCAGCTACTCCAGTAGAGAGGGAGAAACAAAGCAGTAAATGTTGCCACAGCA 660
DB |||||
QY 877 CCACAAAATCTCTGCCAAAGTTGAGGCAACACCCCTTAAAGTTGAGAAGCCATCAGTT 936
DB |||||
QY 661 TCAGATCTCCAAAGTGGATTTGCTACTGATCTGTTTAAACATGCTATCAATGATGAT 720
DB |||||
QY 937 GGGCCGCTCTTAAGTTGATTTATGCTACTGATCTCTTTAAACATGTTATCAATGATGGA 996
DB |||||
QY 721 TCGACTACAAATACCTCAGAGGCACTCTCGGCGATCTCTCGGCGATCAATCACTATCG 780
DB |||||
QY 997 ACAACAGAGAAAGAGTCAGATCATCTCAAA-----CGATGATATGCTGCG 1044
DB |||||
QY 781 GCTGGCTTTTCAGTCTGCTGGAAGTGTCAAAAGGAGAGAAATTTGTCAACGCCAGCT 840
DB |||||
QY 1045 GATGGCTCCAGTCTGCACAAACAGTACCTAGCTCGGAGGAAAGATTTCTGCCAAACCA 1104
DB |||||
QY 841 GCTGAGAGAGATTTCTCTCAGCTTCACTCTCTGA-----CTTTGAGGATTTGTTTAA 894
DB |||||
QY 1105 GCAGAAAGCATGACCCAGTCTACATCTGGAATAGAGACTTATTTAAAGATTTCCGCGATT 1164
DB |||||
QY 895 GACACACCTTAATTTAACTCAACAGCAGCAAAAGATGTCAAGGGCGATATCATGAGC 954
DB |||||
QY 1165 GTGTCTATATCTCAGCTCAGCTGTTTCCCAAGTAAATGCAAGAAATGATCATGAT 1224
DB |||||
QY 955 CTGTTTGAGAGAGCAATATAGTATCGCCTTTTGGCCATGATCATCAGCAACAGGTGCTATG 1014
DB |||||
```

```
DB 1225 TTGTTTGAGAAAGTCCAAACATGCGCATCGCCATTGCTGTCCAAACAGCAGCGCTGGCGTTT 1284
QY 1015 CTCGCTCAGCAGCAGCCCTTTACATGCTGCTCAGCGAAAGCTG 1057
DB 1285 ATGCTCAGCAGCAAGCTCTTCTTAATGGCTGCTCTTAAAGCTG 1327

RESULT 15
US-10-425-115-176513
; Sequence 176513, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176513
; LENGTH: 2221
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2221)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_92572C.1
US-10-425-115-176513

Query Match      23.0%; Score 333.4; DB 20; Length 2221;
Best Local Similarity 58.7%; Pred. No. 1.5e-91;
Matches 624; Conservative 0; Mismatches 421; Indels 18; Gaps 2;

QY 1 ATGACGAGAAAGCCAAAGCTCTCTAAGGAGCTTAATGCCCGCCATAGAAAGATTTCTTGAA 60
DB |||||
QY 261 ATGACGAGAAAGCGTCCGCTCTCCAGAGCTCAAGCCAGCAAGAGATATTGGAA 320
DB |||||
QY 61 GGGCTTTCTTAAACATCCAGAGAAACAGAGATGTGCTGATGCTCAAAACAAAAGTCCAAAG 120
DB |||||
QY 321 GGTCTTCTACGGCATCTCTGAGATAGAGAAATGCGCAGACTGCAAGTCAAAGGTCTCTCGA 380
DB |||||
QY 121 TGGGCTAGTGTTAATTTAGTATCTTTATCTGCATGCAATGTTCTGGGATTCACAGAGT 180
DB |||||
QY 381 TGGGCAAGTGTGAATCTAGGTATCTTTATATGCAATGATGTTCTGGCATTCATAGAAGC 440
DB |||||
QY 181 CTCGGGTTACACATATCGAAGTTCGATCTGCCACTCTGGACACATGGCTCCCGAGCAG 240
DB |||||
QY 441 CTGGGGTGCACATATCTAAGGTAAAGTCTGCTACCTCTGGATACATGCTGCCAGCA 500
DB |||||
QY 241 GTTGCAATTTATACAGTCAATCGGAAATGATAAGCAAAATAGTTACTGGAAGCAGAGCTA 300
DB |||||
QY 501 GTTGCAATTTATCCAATCAATCGGAAACGAAAGCAAAATAGCCATTGGGAAGCAGAGCTG 560
DB |||||
QY 301 CCCCCAACTATGATAGAGTTGGAATTTGAGAAATTTATACGTGCAAGATGTAAGAGAG 360
DB |||||
QY 561 CTTCTAATCTATGATAGGTTGGATAGAGAAATTCATCTCGTGCAGAAATATAGGAGCAAG 620
DB |||||
QY 361 AGATGGGTTCTAGAGGGGAAAGGCTAGATCACTCTCTAGATCGAGCAGGAACGGGG 420
DB |||||
QY 621 AGATGGGTACCGAGAAATGGAACATTTGCGACCTTCTCCGTTTCGAGATGAGAGAGC 680
DB |||||
QY 421 AAATCTGTGGAGAGAGTGGCGGGGATATGAGCATGGAATAGTAGTAGTCTCTTAAT 480
DB |||||
QY 681 CAAGAGTCTCCAGCTAGTGTCTTAATAGGAGTGCACATCATAGATCTTCTGTTGAGCAAAAC 740
DB |||||
QY 481 TTGTTTGAGAGAGGAAACCTATTTCAGCATCTAGACATCTAGAAATATGTTGCTCAAGC 540
DB |||||
```

Db 741 CGTGTCTTACCAGCTCTCCGAGCAAAGTTGCACATGTAGTCTCAAGGATACCCCTCTCAG 800
Qy 541 AGAATAAATCTTCCGTGCTCCCAAGGACCCAGTCAGGTTATAAAGCCACAGCAGAA 600
Db 801 GCATCACCTCAGGCCCCAAAGGTAGAACCCACAGTTCTTAAGGTGGTTTCACTCTCTCAG 860
Qy 601 ATGGAGTCTGCAGCTACTCTCAGTAGAGGGAGAAACAGCAGTAAATGTTGCACCAGCA 660
Db 861 CCACAAAATCTTTCGCAAGTTGAGGCAACACCCCTAAAGTTGAGAAGCCATCAGTT 920
Qy 661 TCAGATCTCCAAAGGTGATTTTGCTACTGATCTGTTTAAACATGCTATCAATGGATGAT 720
Db 921 GCGCGCTCTCTAAAGTTGATTTATGCTACTGATCTCTTTAACAATGTTATCAATGGATGGA 980
Qy 721 TCGACTACAAATACCTCAGAGGCAACTCTCTGGGATACTCTCTGCCGATGATAACTCATGG 780
Db 981 ACAAAGAGAAAGAGTCAGAGTCATCTCAAA-----CGATGATAATGCCCTGG 1028
Qy 781 GCTGGCTTTTCACTGCTGGAAGTGTCAAACGGCAGAGAAATTTGCACAGCCAAAGCCT 840
Db 1029 GATGGCTTCCAGTCTGCACAACCAAGTACCTAGCTCGGAGGAAAGATTTGSCCAACCA 1088
Qy 841 GCTGAGAGCAGTTCTCTCCAGCTTCATCTTCTGA-----CTTTGAGGATTTGTTTAA 894
Db 1089 GCAGAAAGCATGACCCAGTCTACATCTGGATAGAAGACTTATTTAAAGATTCGCCAGTT 1148
Qy 895 GACACACCTAATTTAACTCAACAGCACCAGAAAGATGTGAAGGGCGATATCATGAGC 954
Db 1149 GTGTCAATATCTCAGCTCCAGCTGTTTCCCAAGTAAATGCAAGAAATGATATCATGAGT 1208
Qy 955 CTGTTTGAGAGCAATATAGTATGCGCTTTTGCCATGTCATCAGCAACAGGTTGCTATG 1014
Db 1209 TTGTTTGAGAGTCCAAATATGCAATCACCATTGCTGTCCACAGCAGCAGCTGCGGTTT 1268
Qy 1015 CTCGCTCAGCAGCAAGCCCTTTTACATGGCTGAGCGAAAGCTG 1057
Db 1269 ATGTCTCAGCAGCAAGCTCTTCTAATGGCTGCTCTTAAAGCTG 1311

Search completed: September 27, 2005, 07:23:19
Job time : 3616 secs